Generation of Mismatch Repair Fusion Proteins for Single-Molecule Analysis

Honors Research Thesis

Presented in partial fulfillment of the requirements for graduation with honors research distinction in Biochemistry in the undergraduate colleges of The Ohio State University

by

Jared Bennett

The Ohio State University

April 2014

Project Advisor: Dr. Richard Fishel, Department of Molecular Virology, Immunology, and Medical Genetics

CONTENTS

Overview

Single molecule visualization studies require proteins that are labeled with known fluorophores at known locations on the molecule. This project examines the viability of one method that is capable of specifically and efficiently labeling proteins with fluorophores. The labeling method uses the *F*ormylglycine *G*enerating *E*nzyme (FGE) that recognizes a hexa-amino acid sequence and catalyzes the conversion of a central cysteine into a formyl-glycine containing a reactive aldehyde group. The converted aldehyde may be used in conjunction with Hydrazino-Pictet-Spengler (HIPS) ligation to attach a high quantum yield fluorophore such as Cy3 or Cy5 to a specific site within any protein. We used MutS as an initial test of this method. The fusion cloning method was then extended to the eukaryote PCNA trimer. We then cloned and modified the RecJ protein to more fully test the labeling methodology. RecJ is a 5'-3' exonuclease that functions in conjunction with MutS and MutL to execute bacterial MMR. To date, no exonuclease has been labeled and visualized for single molecule analysis. Finally, in order to test whether fluorophore-labeling RecJ affected its activity, we developed a novel fluorophore-based exonuclease assay.

Introduction

Single molecule total internal reflection fluorescence (smTIRF) microscopy is becoming an important tool for imaging the kinetics, mechanisms and functions of biological complexes (Forties and Wang, 2014). It has the advantage of being direct observation, thus providing detailed information about how individual molecules behave in real-time *in vitro* and *in vivo* (Forties and Wang, 2014). However, there are some technical difficulties associated with smTIRF. Visualization is limited by the diffraction limit of the microscope: ~300 nm for the visible lasers used for TIRF activation in our studies (Meinhart and Wereley, 2003) (figure 1) . Moreover, molecules must be reliably labeled with fluorophores to be visualized (Ha and Tinnefeld, 2012) and labels must be small so that they do not interfere with molecular diffusion rates or interactions (Monico et al., 2013). A common method for fluorophore labeling has been accomplished by fusing a green fluorescent protein (GFP) or its color derivatives to proteins of interest. However, GFP tags are large (23-25 kD or >200 amino acid residues), are relatively low quantum yield (number of times a specific radiation event occurs per photon absorbed), and easily photo-bleached (Axelrod et al., 1976). Quantum dots and chemical fluorophores have a robust quantum yield, but they must be specifically attached to the protein by often toxic chemical processes (Jaiswal and Simon, 2004). Moreover, quantum dots are on the order of 100 nm in diameter, making their size an issue in some single molecule studies (Medintz et al., 2005). For the relatively small, high quantum yield chemical fluorophores, traditional chemistry attaches *N*-Hydroxysuccinimide-derived fluorophores to exposed cysteine residues within proteins using maleimide chemistry (Sharpless and Flavin, 1966). This is effective in small proteins with one exposed cysteine, but not in most proteins where there are generally multiple cysteines, including cysteines required for structure or catalysis.

Here, we describe the further development of a unique hexa-amino acid peptide tag that can be selectively converted into a formylglycine containing a distinct reactive aldehyde group (Carrico et al., 2007). *F*ormyglycine *G*enerating *E*nzyme (FGE) recognizes this unique hexa-amino acid tag and converts the central cysteine into a formylglycine. A hydrazide-modified fluorophore may then react with the aldehyde via hydrazone chemistry resulting in covalent attachment (figure 3). Unfortunately the hydrazide-hydrazone linkage is highly reversible (Christie et al.). Thus, in collaboration with Redwood Biosciences (Emeryville, CA), we have developed a novel fluorophore linker that is based on

the *H*ydraz*i*no-*P*ictet-*S*pengler (HIPS) ligation (Agarwal et al.) (figure 2). These HIPS-flourophore reagentsdisplay high specificity for a reactive aldehyde and are completely irreversible. The HIPS-flourophore technology allows for specific labeling of FGE-tagged proteins that have been converted by the FGE enzyme at only one site.

While all genomes are under constant threat from exogenous damaging factors (Ciccia and Elledge, 2011), one of the most important potential mutation-causing circumstances occurs during DNA replication (Tippin et al., 2004). The DNA polymerase machinery has an intrinsic nucleotide misincorporation rate of ~1 nucleotide per 10^6 nucleotides copied (Tippin et al., 2004). Mismatch repair (MMR) is a highly conserved replication-coupled DNA repair system that has been shown to reduce this error rate approximately 1000-fold (Kolodner and Marsischky, 1999; Modrich, 1997). MMR is a bi-directional excision-resynthesis reaction where mismatch recognition is transmitted to a distant strand scission either 3' or 5' of the mismatch (Kolodner and Marsischky, 1999; Modrich, 1989). Excision of the error-containing strand begins at the strand scission and continues to just past the mismatch (Lahue et al., 1987). The importance of MMR can be seen in individuals that inherit a single MMR gene defect, which causes the most common cancer predisposition syndrome Lynch syndrome or hereditary non-polyposis colorectal cancer (LS/HNPCC). Other MMR-dependent cancer syndromes (MMRCS) include Turcot syndrome and Constitutive Mismatch Repair Deficiency (CMMR-D), where compound heterozygous or homozygous MMR mutation carriers present with very early and aggressive childhood cancers (Martin-Lopez et al.). Classical LS/HNPCC is largely caused by mutations in hMSH2 and hMLH1 (Martin-Lopez et al.). In addition to hMSH2 and hMLH1, MMRCS may include mutations in hMSH6 and hPMS2 (Martin-Lopez et al.).

The hMSH2-hMSH6 heterodimeric proteins are the human homologs of the bacterial MutS. MutS functions as a homodimer that scans the DNA for mismatched or damaged nucleotides (Su and Modrich, 1986). It then recruits dimeric MutL to aid in repair (Grilley et al., 1989). The hMLH1-hPMS2 heterodimeric proteins are the human homologs of bacterial MutL. In gram-negative enteric bacteria the error-containing strand following nucleotide misincorporation during replication is determined by the MutH protein, which introduces a strand-specific scission into a nearby transiently unmethylated *D*NA *a*denine *m*ethylase (Dam) site (G↓ATC/G^{me}ATC; (Marinus, 1976; Welsh et al., 1987). This distant strand scission is the initiation site for the exonuclease excision reaction (Lahue et

al., 1987). In *E.coli*, four redundant exonucleases, Exo1 and Exo 10 for 3'-5' excission and Exo 7 and RecJ for 5'-3' excission, have been identified as excision-exonucleases for MMR (Viswanathan and Lovett, 1998). RecJ is one of the two 5'→3' MMR exonucleases (Sutera et al., 1999). It appears to function as one of the bacterial equivalents of hExoI that is required for human MMR.

The Proliferating Cell Nuclear Antigen (PCNA) functions as a heterotimeric protein sliding clamp (Bowman et al., 2004). It enhances replicative polymerase processivity and links the MMR machinery to replication (Lau and Kolodner, 2003). The function of PCNA in replication and MMR is highly conserved and in bacteria is replaced by the β-clamp (Lau and Kolodner, 2003). This highly conserved nature of the MMR system implies that studies with the simplified bacterial system will allow us to dissect protein functions and ultimately apply this knowledge to the study of human MMR. By understanding MMR protein activities it may be possible to determine how to alleviate or bypass the functional defects that lead to LS/HNPCC and MMRCS.

The detailed mechanism of MMR is enigmatic. There are several models that have been proposed for the individual and complex functions of the MMR proteins (Acharya et al., 2003; Fishel et al., 2000; Kolodner et al., 2007). Unfortunately, previous biochemical studies do not unequivocally support any one MMR mechanism. Through the use of fluorescently labeled proteins and real-time single molecule analysis it is likely that the complete MMR mechanism may be visualized and ultimately deduced. This honors project is part of a larger study designed to understand the biophysical mechanics of MMR. A method for specific labeling of proteins with chemical fluorophores will be tested and determined to be effective. That method will then be expanded to other proteins necessary to the study of MMR. Every protein component of MMR must be examined by itself, in conjunction with well-defined MMR components, and within the whole MMR process to elucidate the basic mechanisms underlying MMR.

Total Internal Reflective Fluorescence (TIRF)

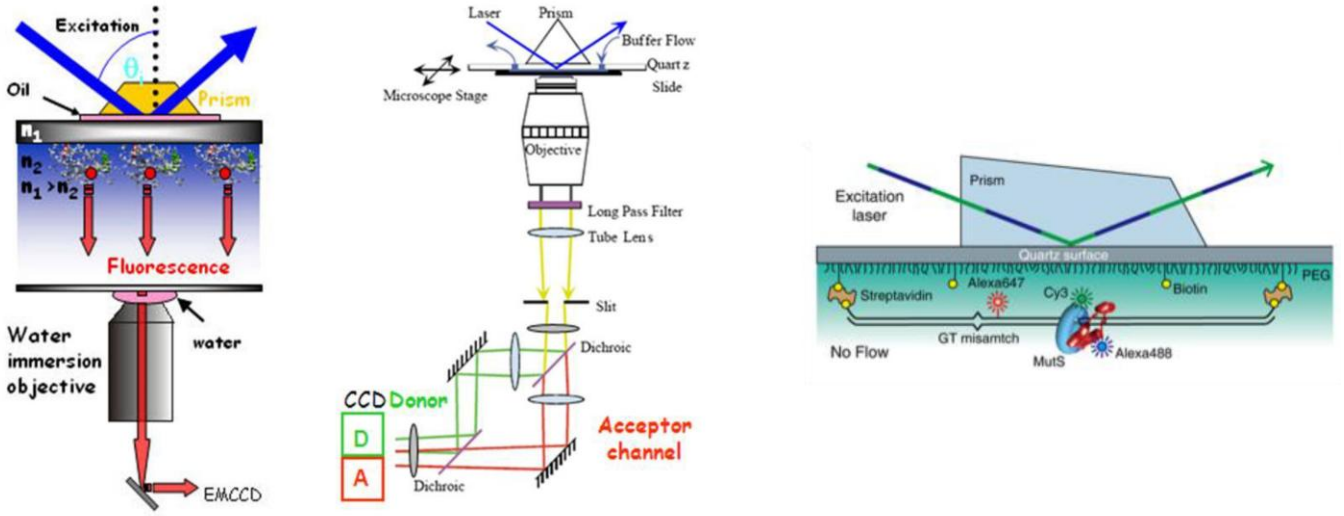

Figure 1. TIRF and Single-Molecule Study Designs. On the left is an outline of how a Total Internal Reflective Fluorescence (TIRF) experiment is set up. A laser is shown through a prism at a critical angle, θ. This causes an effervescent wave to propagate up to 100 nm into a sample. This excites fluorophores close to the surface. The light emitted from these fluorophores is captured by a camera filtered to only find light of that specific wavelength. On the right is a Single-Molecule design used to study interactions on DNA. The DNA is bound close to the surface, and proteins can diffuse along it. The problem with this setup is that it is diffraction limited, $d = \lambda / [2*n*\mathrm{Sin}(\theta)]$, where d is the radius of resolution., With our setup, we have a resolution of about 300 nm. Also, every protein and DNA molecule must be labeled with a fluorophore to be visualized.

Hydrazino-Pictet-Spengler (HIPS) Dye

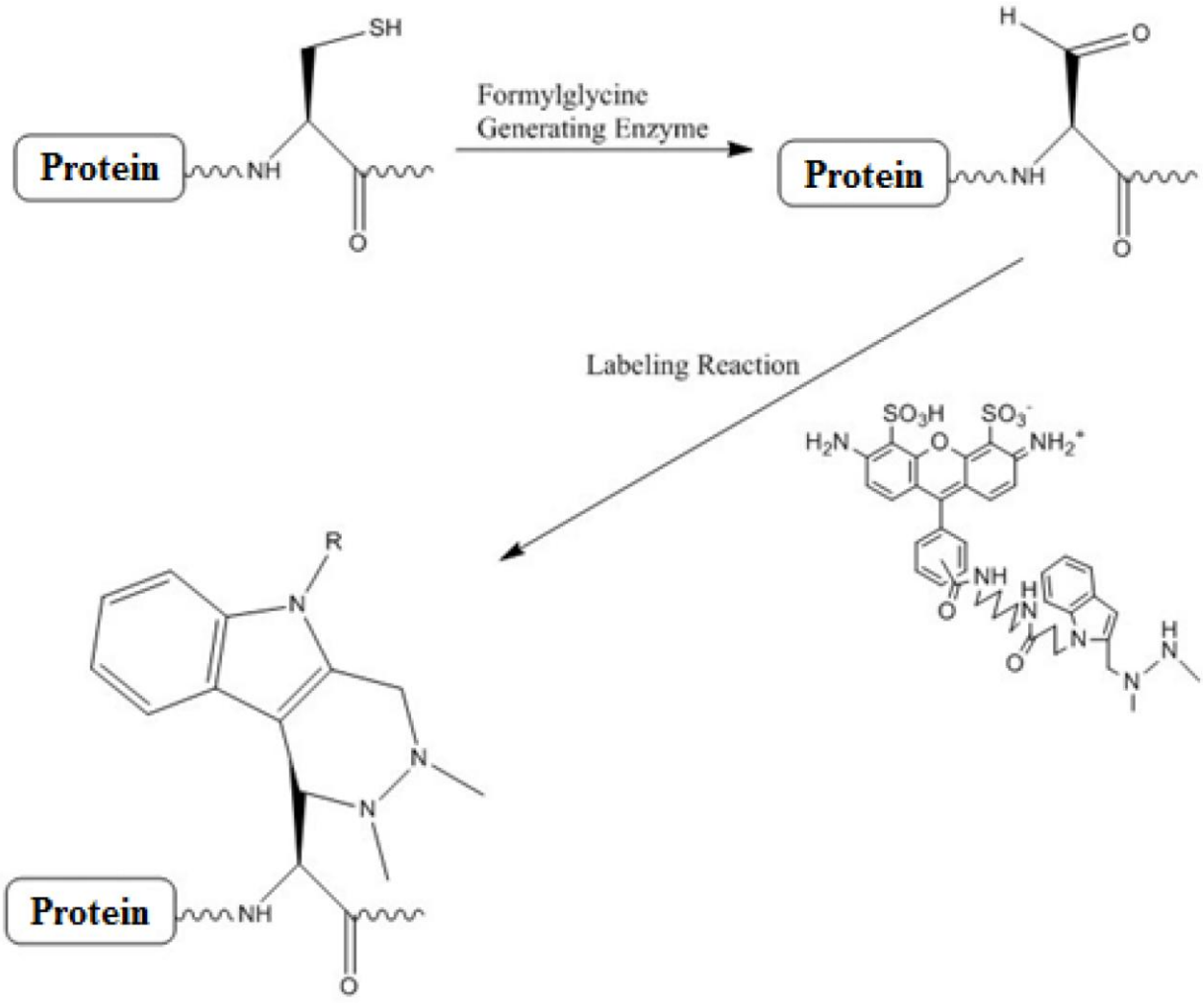

Figure 2. HIPS Dye Reaction. After conversion of the cysteine into a formylglycine by FGE, the HIPS dyes react with the formylglycine to form an irreversible bond.

Hydrazone-Hydrazine Chemistry

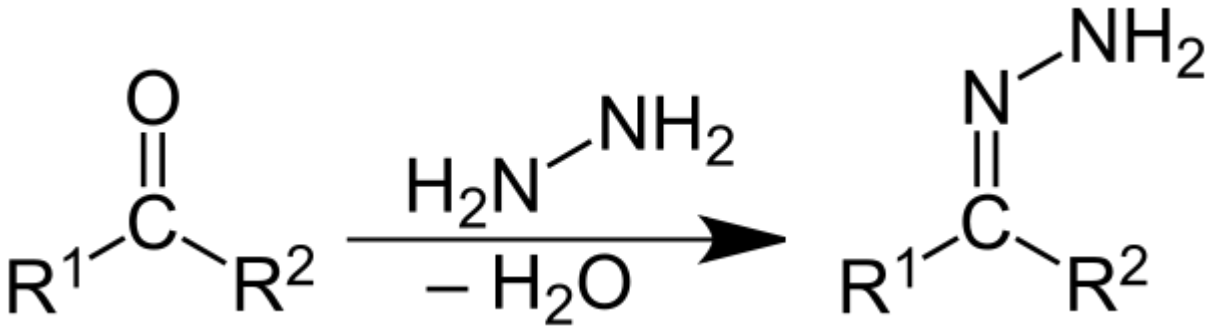

Figure 3. Hydrazone-Hydrazine Chemistry. The formylglycine generated by FGE will also react to form a hydrizine. This was the original chemistry used to label FGE products. However, it is highly reversible, and not suitable for Single-Molecule analysis.

Chapter 1

Construction of Mismatch Repair Fusion Proteins

Introduction:

MutS is a sliding clamp that is part of the MisMatch Repair (MMR) process in bacteria (Lahue et al., 1989). It functions alongside MutL, MutH, UvrD, and an exonuclease (Viswanathan and Lovett, 1998). MutS recognizes nucleotide mismatches in DNA (Su and Modrich, 1986). It then forms a stable clamp around the mismatch by binding ATP (Acharya et al., 2003). Impediment of ATP binding or hydrolysis will arrest MutS function at intermediate steps in this reaction (Cho et al., 2012). This is the basis for generating two MutS mutants: MutS(K620A) and MutS(K620R). MutS(K620A) will not bind ATP, thus inhibiting all further steps (Jeong et al., 2011). MutS(K620R) binds but does not hydrolyze ATP, thus preventing release of MutS (Jeong et al., 2011).

Proliferating Cell Nuclear Antigen (PCNA) is a sliding clamp that increases the processivity of replicative polymerases (Jonsson et al., 1998). It also localizes accessory factors to sites of replication. Replication is a potential source of single-base mutations (Tippin et al., 2004). As such, MMR machinery is localized with the replication fork to act as a replication-coupled DNA repair mechanism (Lau and Kolodner, 2003). PCNA is necessary to extend MMR studies to a human system. Labeling with a fluorophore will allow for bulk Fluorescence Resonance Energy Transfer (FRET) with other members of the human MMR system.

RecJ is a 5' to 3' non-processive exonuclease (Sutera et al., 1999). It acts to recise 5' overhangs generated during MMR recombination via recBCD-independent recombination (Han et al., 2006). RecJ is necessary for the production of 3' ssDNA overhangs used for strand-invasion after a double stranded break (Viswanathan and Lovett, 1998). An exonuclease is an essential part of the MMR reaction *in vitro*. As such, it is necessary to purify and label it for single-molecule analysis of the MMR excision mechanism.

Methods:

Materials **-** Ultrapure chemicals were purchased from Amresco (Solon, OH) and Sigma Aldrich (St. Louis, MO). Oligonucleotides were designed and purchased from IDT (Coralville, IA). MutS was retrieved from a laboratory stock in pET29a and contained C-terminal (FGE) and His_6 tags. Wild-type PCNA was retrieved from a laboratory stock pET23b. Wild-type RecJ was gratefully received from the laboratory of Dr. Susan Lovett (Bandies University) in STL10 pRDK115 and STL1739 pWSK29-CJS1 vectors containing the b-lactamase (Amp^r) gene.

Point mutations and fusion tags **-** Point mutations of MutS were generated using the QuikChange Site-Directed Mutagenesis according to the manufacturer's recommendations (Stratagene). Fusion tags were generated using the same scheme outlined in the QuikChange Site-Directed Mutagenesis protocol but with external unannealed DNA flaps (Figure 1). The length of fusion tags required a two-primer design to implement. PCNA was initially amplified using an inner set of primers (Figure 8). The reactions were done in bulk before being divided into several lanes for gel purification using 1% agarose gels (figure 8). The gels were visualized using Ethidium Bromide (0.5 ug/ml) under short wave UV fluorescent illumination . These lanes were excised and purified as one prior to being used as substrate in a second reaction. Gel extraction of DNA products was performed with the Qiagen gel-extraction kits according to manufacturer recommendations (Qiagen, Germantown, MD). The second reaction amplified PCNA, using outer primers, to generate the full fusion tags (Figure 8). The second reaction was also done in bulk, then divided amongst several lanes for gel purification. These reactions were then excised and purified as one product prior to digestion and ligation into plasmidsPlasmid and DNA digestions were done using New England Biolabs (NEB) enzymes and conditions (Beverly, MA).

Construction of fusion derivatives **-** PCNA was digested with Nde I on the N-terminus and BamHI on the C-terminus prior to being ligated into pET3a and pET9a. RecJ was digested with Xba I on the N-terminus and BamHI on the C-terminus prior to being ligated into pET3a and pET9a. Ligations were done with NEB DNA T4 ligase at 16^{o}C overnight as suggested by the manufacturer (NEB, Beverly, MA). Plasmids were transformed into XL-10 Gold UltraCompetent cells from Agilent Technologies (Santa Clara, CA) using a heat-shock technique (30min on ice, 45sec at 40^{o}C, 2min on ice) followed by 30 min shaking growth with 100uL of LB (Miller's Recipe, EMD Millipore). Cellular DNA extraction

was performed using Qiagen miniprep kits according to the manufacturer protocol (Qiagen, Germantown, MD), Products were verified via sequencing, using the defined primers (Figures 1, 4, and 9). Primer allocation, for complete gene coverage is shown in Figures 2, 5, and 10.

Results:

Construction of site-specific MutS mutation in the ATP binding domain - MutS was retrieved from a laboratory stock already containing a C-terminal FGE-His_6 fusion tag. This project was concerned with creating two substitution mutant forms: an ATP-binding deficient form [MutS(K620A)], and an ATP-binding proficient but ATP-hydrolysis deficient form [MutS(K620R)]. We designed oligonucleotide primers that would hybridize to the DNA surrounding the sequence that encodes the ATP-binding P-loop domain (Walker et al., 1982) ;Figure 1). For the MutS(K620A) substitution, we altered three coding nucleotides from AAA→GCG in the forward and reverse primers in the QuikChange Site-Directed Mutagenesis protocol (Figure 1). For the MutS(K620R) substitution, we altered three coding nucleotides from AAA→CGT in the forward and reverse primers in the QuikChange Site-Directed Mutagenesis protocol (Figure 1). The overall strategy is shown in Figure 3. The PCR products were isolated and transformed into XL10 Gold and the entire gene and surrounding promoter region of several independent transformants were sequenced (Figure 1 and 2). Two subclones that contained the desired mutations in the ATP binding domain but were otherwise *wild type* for the MutS coding and expression sequences were identified.

Construction of PCNA fusion derivatives - A wild type PCNA clone in the pET23b expression vector was retrieved from a laboratory stock. This project was to create several fusion constructs to be used for purification and fluorophore labeling studies. Primer designs for the fusion constructs and gene sequencing are shown in Figures 4 and 5. The overall strategy is shown in Figures 6 and 7. Fusion tags were added to the N-terminus of PCNA because the C-terminus is involved in protein-protein interactions necessary to its function in replication. The first construct was a simple N-terminal His_6-PCNA fusion that could be used for large-scale purification of the largely wild type PCNA protein using a Nickel(Ni)-NTA metal affinity chromatography. Ni-NTA is specifically coordinated by the hexa-His tag allowing a substantial enrichment of His_6-tagged proteins from contaminating proteins during purification (Bornhorst and Falke, 2010). This construct is the only one not to require the dual-primer method to generate a complete fusion tag. The His_6 tag is short enough to only require one

primer. The second construct was a N-terminal His_6-HRV-FGE-PCNA fusion containing a HRV-3C 8-amino acid protease recognition site located between the His_6-tag and the FGE-tag. The location of the HRV-3C site will allow Ni-NTA purification followed by proteolytic cleavage and release of the His_6 site leaving the FGE-tag that can be used in fluorophore labeling. Aside from removing none native amino acids, the advantage of the HRV-3C protease is that it functions efficiently at 4^oC where most isolated proteins may resist thermal denaturation and inactivation. Because of the length of these tags, compared to the simple His_6 tag, two primers were necessary for efficient PCR amplification. This construct, and all remaining constructs, used two primers to fully generate the fusion tags.

The third and fourth constructs contain a N-terminal FGE-His_6-PCNA fusion and N-terminal FGE(mut)-His_6-PCNA fusion. Recent laboratory observations have suggested that the central cysteine of an internal FGE hexa-amino acid sequence is not converted as efficiently as an external FGE hexa-amino acid sequence. These observations led to the design of primers for the construction of a N-terminal FGE-His_6-PCNA fusion (Figure 4). To control for the specificity of the FGE labeling we also constructed a mutant version of the FGE hexa-amino acid recognition site that replaced the central cysteine with an alanine residue [FGE(mut)] (Figure13). The overall strategy for the PCR of fragments and subcloning into the PCNA expression plasmid are shown in Figure 7. All of these constructs were cloned into pET3a and pET9a so they could be expressed from plasmids containing either ampicillin resistance or kanamycin resistance, respectively.

There are four different PCNA constructs placed into two vectors, pET9a and pET3a, for a total of eight different constructs (Figures 6 and 7). The pET9a PCNA constructs were successfully completed and verified by sequencing. The pET3a constructs were also completed successfully.

Construction of RecJ fusion derivative **–** The RecJ wild-type gene a gift from Dr. Susan Lovett (Brandeis Unversity). This project was concerned with attaching His_6 and FGE tags to the C-terminus, via the same dual-primer design used for PCNA fusion tags. The C-terminus was chosen based upon crystal structures of RecJ. The crystal structures had the N-terminus and C-terminus removed because of the difficulty in crystallizing them. This implied that they were both unstructured. It was determined later, through labeling studies, that the C-terminus blocks access by FGE to the tag. Another set of fusion derivatives with N-terminal tags was generated by another member of the lab.

These tags allow for simplified purification and labeling procedures for single-molecule studies. These constructs were also put into pET3a and pET9a.

The RecJ construct was successfully cloned into two vectors, pET9a and pET3a (Figure 12). The pET3a construct has mutations within the RecJ coding sequence that were in the process of being corrected when issues with labeling became apparent (see Chapter 2). The pET9a constructs were successful and the RecJ coding sequences was verified as intact by sequencing.

Discussion:

MutS **-** MutS mutants were successfully generated allowing for the study of MutS on DNA when it can not bind ATP, or when it can bind but can not hydrolyze ATP. The mechanism of MutS function can now be broken into several discrete steps to be analyzed. Subsequent purification and labeling with defined fluorophore(s) are necessary prior to its use in single-molecule TIRF studies.

PCNA **-** Eight different PCNA constructs were created. Sequencing data shows that the specific tags were introduce in-frame with an intact PCNA gene. Each construct may be used in different biochemical analysis in the MMR reaction, for labeling specificity, or as simple His-tagged proteins as a co-factor in other experiments. Fluorophore labeled constructs can be used in Fluorescence Resonance Energy Transfer (FRET) experiments with other members of the MMR machinery: hMSH2-hMSH6, hMSH2-hMSh3, hMlh1-hPMS2, ExoI among others. Further expression studies will be required to allow for production and purification of PCNA for use in these experiments.

RecJ **-** RecJ constructs generated by PCR amplification were analyzed by agarose gel electrophoresis. Construct integrity and proper fusion tags were confirmed via sequencing. Since fusion tags are an unnatural addition to RecJ, *in vivo* complementarity must be tested to insure proper function has not been inhibited. The His_6 tag will aid in purification of RecJ to be used in labeling experiments, using the FGE tag, and kinetics experiments in a novel, fluorophore based assay.

MutS Mutant Designs

Mutagenesis Primers		
K → A Forward	ccggtccgaacatGGGCGGTGCGAGTACCTATATGCGCC	Tm= 74 C
K → A Reverse Complement	GGCGCATATAGGTACTCGCACCGCCCatgttcggaccgg	Tm= 74 C
K → R Forward	ccggtccgaacatgggcggtcgTagtacctatatgcgcc	Tm= 74 C
K → R Reverse Complement	ggcgcatataggtactAcgaccgcccatgttcggaccgg	Tm= 74 C
Sequencing Primers		
muts seq f1-cc	gat cct gct gtt tta	
muts seq F2-cc	tcg aca ccg cgc gcc	
Muts seq F3-CC	gat agt gca ccg gta	
Muts seq F4-cc	gaa cta cac ctg ccc	
muts Seq nterm-CC (RC)	tccggcgaccagcaa	
muts seq cterm-Cc	tgc atc tcg atg cac	
MUTS SEQ F1-Reverse-JB	gggacccgcgatctg	

Figure 1. MutS Primer Design. MutS mutants, lysine 620 to alanine and lysine 620 to arginine, were generated using the first four primers listed. The primers were designed to cover the exact same amino acid and amplify the plasmid in opposite directions, thus introducing a single amino acid change into both strands. This design creates an overlapping nick that allows for the plasmid to anneal and be sealed *in vivo.* Beneath that are primers for sequencing

MutS 187 kD Dimer - Primer Allocation

ATGAGTGCAATAGAAAATTTCGACGCCCATACGCCCATGATGCAGCAGTATCTCAGGCTGAAAGCC
CAGCATCCCGAGATCCTGCTGTTTTACCGGATGGGTGATTTTTATGAACTGTTTTATGACGACGCAAA
ACGCGCGTCGCAACTGCTGGATATTTCACTGACCAAACGCGGTGCTTCGGCGGGAGAGCCGATCCC
GATGGCGGGGATTCCCTACCATGCGGTGGAAAACTATCTCGCCAAACTGGTGAATCAGGGAGAGTC
CGTTGCCATCTGCGAACAAATTGGCGATCCGGCGACCAGCAAAGGTCCGGTTGAGCGCAAAGTTGT
GCGTATCGTTACGCCAGGCACCATCAGCGATGAAGCCCTGTTGCAGGAGCGTCAGGACAACCTGCT
GGCGGCTATCTGGCAGGACAGCAAAGGTTTCGGCTACGCGACGCTGGATATCAGTTCCGGGCGTTT
TCGCCTGAGCGAACCGGCTGACCGCGAAACGATGGCGGCAGAACTGCAACGCACTAATCCTGCGG
AACTGCTGTATGCAGAAGATTTTGCTGAAATGTCGTTAATTGAAGGCCGTCGCGGCCTGCGCCGTCG
CCCGCTGTGGGAGTTTGAAATCGACACCGCGCGCCAGCAGTTGAATCTGCAATTTGGGACCCGCGA
TCTGGTCGGTTTTGGCGTCGAGAACGCGCCGCGCGGACTTTGTGCTGCCGGTTGTCTGTTGCAGTAT
GCGAAAGATACCCAACGTACGACTCTGCCGCATATTCGTTCCATCACCATGGAACGTGAGCAGGAC
AGCATCATTATGGATGCCGCGACGCGTCGTAATCTGGAAATCACCCAGAACCTGGCGGGTGGTGCG
GAAAATACGCTGGCTTCTGTGCTCGACTGCACCGTCACGCCGATGGGCAGCCGTATGCTGAAACGC
TGGCTGCATATGCCAGTGCGCGATACCCGCGTGTTGCTTGAGCGCCAGCAAACTATTGGCGCATTGC
AGGATTTCACCGCCGGGCTACAGCCGGTACTGCGTCAGGTCGGCGACCTGGAACGTATTCTGGCAC
GTCTGGCTTTACGAACTGCTCGCCCACGCGATCTGGCCCGTATGCGCCACGCTTTCCAGCAACTGCC
GGAGCTGCGTGCGCAGTTAGAAACTGTCGATAGTGCACCGGTACAGGCGCTACGTGAGAAGATGG
GCGAGTTTGCCGAGCTGCGCGATCTGCTGGAGCGAGCAATCATCGACACACCGCCGGTGCTGGTAC
GCGACGGTGGTGTTATCGCATCGGGCTATAACGAAGAGCTGGATGAGTGGCGCGCGCTGGCTGACG
GCGCGACCGATTATCTGGAGCGTCTGGAAGTCCGCGAGCGTGAACGTACCGGCCTGGACACGCTGA
AAGTTGGCTTTAATGCGGTGCACGGCTACTACATTCAAATCAGCCGTGGGCAAAGCCATCTGGCAC
CCATCAACTACATGCGTCGCCAGACGCTGAAAAACGCCGAGCGCTACATCATTCCAGAGCTAAAAG
AGTACGAAGATAAAGTTCTCACCTCAAAAGGCAAAGCACTGGCACTGGAAAAACAGCTTTATGAA
GAGCTGTTCGACCTGCTGTTGCCGCATCTGGAAGCGTTGCAACAGAGCGCGAGCGCGCTGGCGGA
ACTCGACGTGCTGGTTAACCTGGCGGAACGGGCCTATACCCTGAACTACACCTGCCCGACCTTCATT
GATAAACCGGGCATTCGCATTACCGAAGGTCGCCATCCGGTAGTTGAACAAGTACTGAATGAGCCAT
TTATCGCCAACCCGCTGAATCTGTCGCCGCAGCGCCGCATGTTGATCATCACCGGTCCGAACATGGG
CGGTAAAAGTACCTATATGCGCCAGACCGCACTGATTGCGCTGATGGCCTACATCGGCAGCTATGTA
CCGGCACAAAAAGTCGAGATTGGACCTATCGATCGCATCTTTACCCGCGTAGGCGCGGCAGATGAC
CTGGCGTCCGGGCGCTCAACCTTTATGGTGGAGATGACTGAAACCGCCAATATTTTACATAACGCCA
CCGAATACAGTCTGGTGTTAATGGATGAGATCGGGCGTGGAACGTCCACCTACGATGGTCTGTCGCT
GGCGTGGGCGTGCGCGGAAAATCTGGCGAATAAGATTAAGGCATTGACGTTATTTGCTACCCACTAT
TTCGAGCTGACCCAGTTACCGGAGAAAATGGAAGGCGTCGCTAACGTGCATCTCGATGCACTGGAG
CACGGCGACACCATTGCCTTTATGCACAGCGTGCAGGATGGCGCGGCGAGCAAAAGCTACGGCCTG
GCGGTTGCAGCTCTGGCAGGCGTGCCAAAAGAGGTTATTAAGCGCGCACGGCAAAAGCTGCGTGA
GCTGGAAAGCATTTCGCCGAACGCCGCCGCTACGCAAGTGGATGGTACGCAAATGTCTTTGCTGTC
AGTACCAGAAGAAACTTCGCCTGCGGTCGAAGCTCTGGAAAATCTTGATCCGGATTCACTCACCCC
GCGTCAGGCGCTGGAGTGGATTTATCGCTTGAAGAGCCTGGTGTAA

Figure 2. MutS Primer Allocation. This sequence of MutS shows where the sequencing primers from Figure 1 are arranged within the gene. This placement allows for complete coverage of the gene during sequencing.

MutS Construct Design

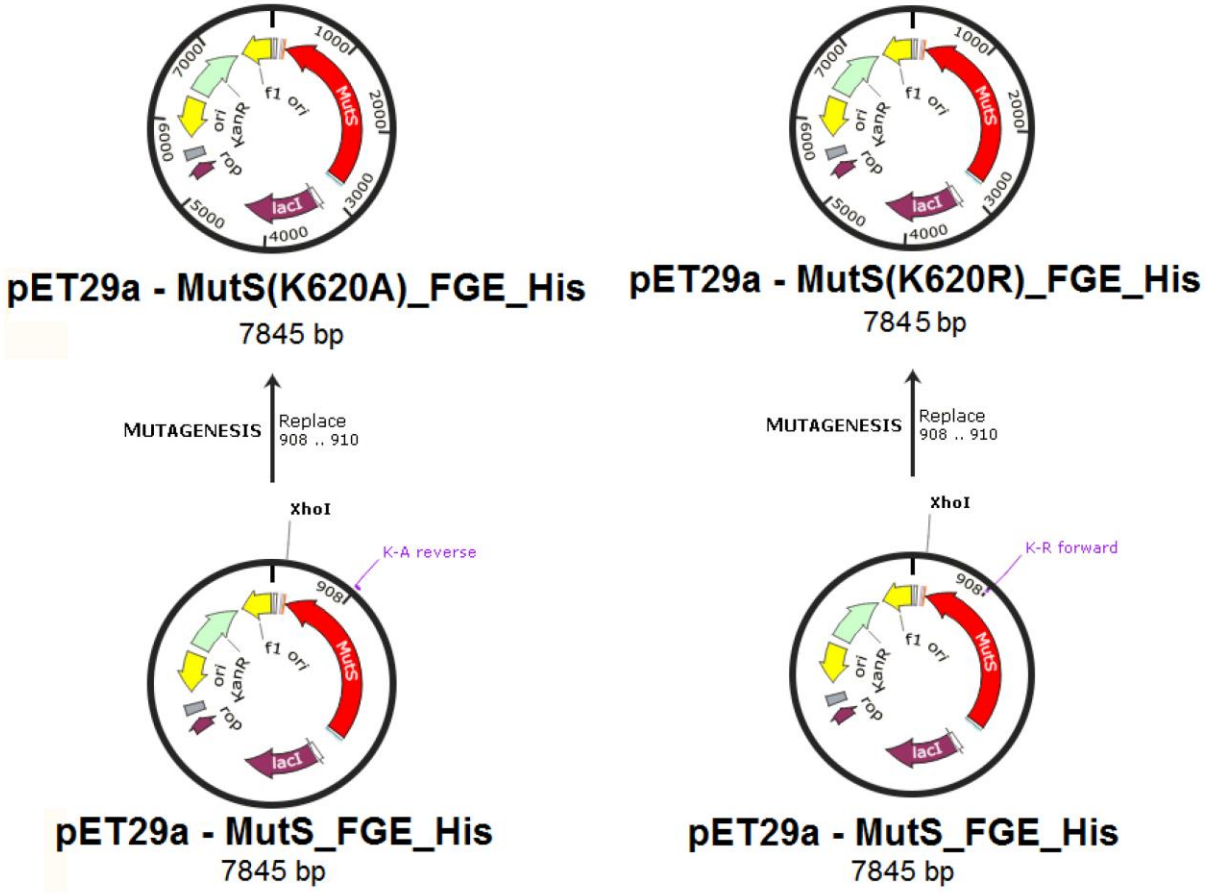

Figure 3. MutS Construct Design. MutS was received in pET29a with C-terminal FGE and His_6 tags already attached. Mutants were designed to retain this previous work while modifying the gene. Thus, primers were designed to cover the nucleotides to be modified and amplify the entire plasmid.

PCNA Fusion Tag Design

Mutagenesis Primers		
+6-NDE1-HIS-Protease-gly- K – gly - FGE-gly-gene-BamHI-+6	N-terminus His_6, protease, and FGE	
HPF Forward 1	gcgcgc - CATATG - cat cat cat cat cat cat - ctg gaa gtg ctg ttt - CAGGGCCCGGGC	Tm= 64 C
HPF Forward 2	cag ggc ccg - ggc - aaa – ggc - ctg tgc acc ccg agc cgt - ggc-ATGTTCGAGGCGCGC	Tm= 64 C
Reverse	GATCGAGGATGAAGAAGGATCTTAG– gga tcc - gcgcgc	Tm= 62 C
+6-Nde1-k - gly- fge – gly – his – gly – gene – BamHI - +6	N-terminus FGE and His_6	
FH Forward 1	gagaga – CATATG - aaa – ggc - ctg tgc acc ccg agc cgt – GGCCATCATCATCATCATCATG	Tm= 62 C
FH Forward 2	ggc - cat cat cat cat cat cat – ggc -ATGTTCGAGGCGCGC	Tm= 64 C
+6-Nde1-k - gly- fge(mut) – gly – his – gly – gene – BamHI - +6	N-terminus FGE (Mut) and His_6	
FmH Forward 1	gagaga – CATATG - aaa – ggc - ctg gcg acc ccg agc cgt – GGCCATCATCATCATCATCATG	Tm= 62 C
+6-NDE1-HIS_6-gly-gene-BamHI-+6	N-terminus His_6	
H Forward	gcgcgc - CATATG - cat cat cat cat cat cat – ggc - ATGTTCGAGGCGCGC	Tm= 64 C
Sequencing Primers		
PCNA Seq 1 F	5'CCT ATA AAT ATT CCG GAT TA 3'	
PCNA Seq 2 F	5'CTA TGA AAT GAA GTT GAT GG 3'	
PCNA Seq 3 R	5'CTC TAG TAC TTC TCG ACA AG 3'	
PCNA Seq 4R	5'CAG CTG TAC TCC TGT TCT GG 3'	

Figure 4. PCNA Primer Constructs. Eight total PCNA constructs were generated. The same cloning sites were used in pET3a and pET9a. Primer designs for the constructs are shown above. Only the His_6 sequence was short enough to be added with one primer. The remaining sequences had the tag divided over 2 primers that were sequentially applied via PCR. Beneath that are sequencing primers used to verify completed constructs.

PCNA 29 KD Hetero-Trimer - Primer Allocation

CCTATAAATATTCCGGATTA(1)TTCATACCGTCCCACCATCGGGCGCGGATCC**GCCACCAT
GCACCACCACCACCACCACGGTGGCGGTCTGAACGACATCTTCGAGGCTCAGAAAAT
CGAATGGCACGAATAAATG**TTCGAGGCGCGCCTGGTCCAGGGCTCCATCCTCAAGAAGG
TGTTGGAGGCACTCAAGGACCTCATCAACGAGGCCTGCTGGGATATTAGCTCCAGCGGTGT
AAACCTGCAGAGCATGGACTCGTCCCACGTCTCTTTGGTGCAGCTCACCCTGCGGTCTGAG
GGCTTCGACACCTACCGCTGCGACCGCAACCTGGCCATGGGCGTGAACCTCACCAGTATGT
CCAAAATACTAAAATGCGCCGGCAATGAAGATATCATTACACTAAGGGCCGAAGATAACGC
GGATACCTTGGCGCTAGTATTTGAAGCACCAAACCAGGAGAAAGTTTCAGA**CTATGAAAT
GAAGTTGATGG(2)**ATTTAGATGTTGAACAACTTGGAATT**CCAGAACAGGAGTACAGCTG(
4)**TGTAGTAAAGATGCCTTCTGGTGAATTTGCACGTATATGCCGAGATCTCAGCCATATTGGA
GATGCTGTTGTAATTTCCTGTGCAAAAGACGGAGTGAAATTTTCTGCAAGTGGAGAACTTG
GAAATGGAAACATTAAATTGTCACAGACAAGTAATGTCGATAAAGAGGAGGAAGCTGTTA
CCATAGAGATGAATGAACCAGTTCAACTAACTTTTGCACTGAGGTACCTGAACTTCTTTAC
AAAAGCCACTCCACTCTCTTCAACGGTGACACTCAGTATGTCTGCAGATGTACCCCTTGTT
GTAGAGTATAAAATTGCGGATATGGGACACTTAAAATACTACTTGGCTCCCAAGATCGAGGA
TGAAGAAGGATCT**TAGGCGGCCGC**TTTCGAATCTAGAGCCTGCAGTCTCGAGGCATGCGG
TACCAAG**CTTGTCGAGAAGTACTAGAG(3)**GATCATAATCAGCCATACCACATTTGTAGAGG
TTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGC
AATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCAC
AAATTTCACAAATAAAG

Figure 5. PCNA Primer Allocation. This PCNA sequence shows where sequencing primers were placed so that sequencing covered the complete gene. These primers were used to verify that fusion tags were correctly implemented and extraneous mutations were not generated.

PCNA Construct Design

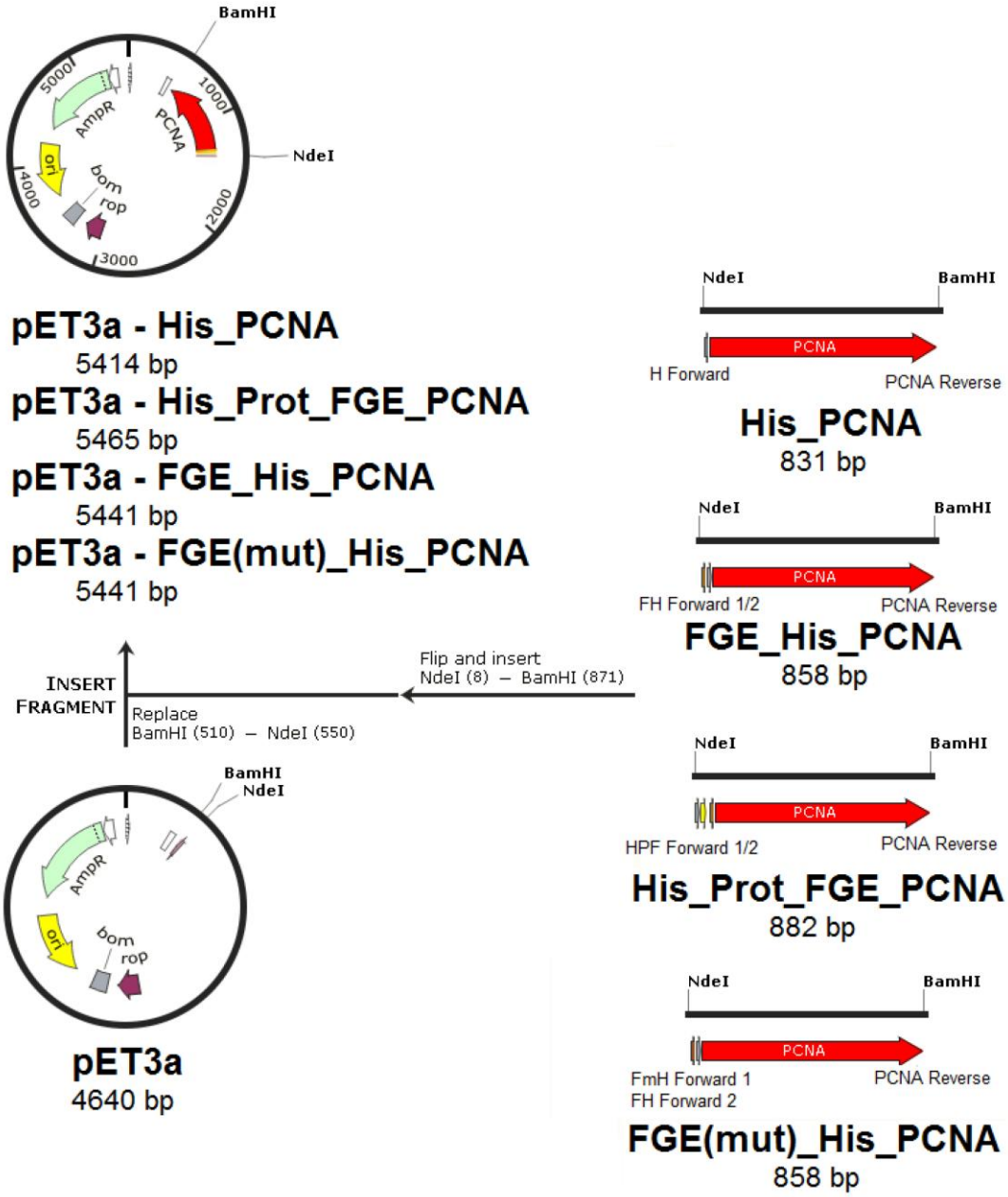

Figure 6. pET3a PCNA Constructs. PCNA was cloned into pET3a (Amp^r) at BamH1 and Nde1 sites. My four constructs are shown on the right. The final constructs are shown at top. Constructs were verified via sequencing.

PCNA Construct Design

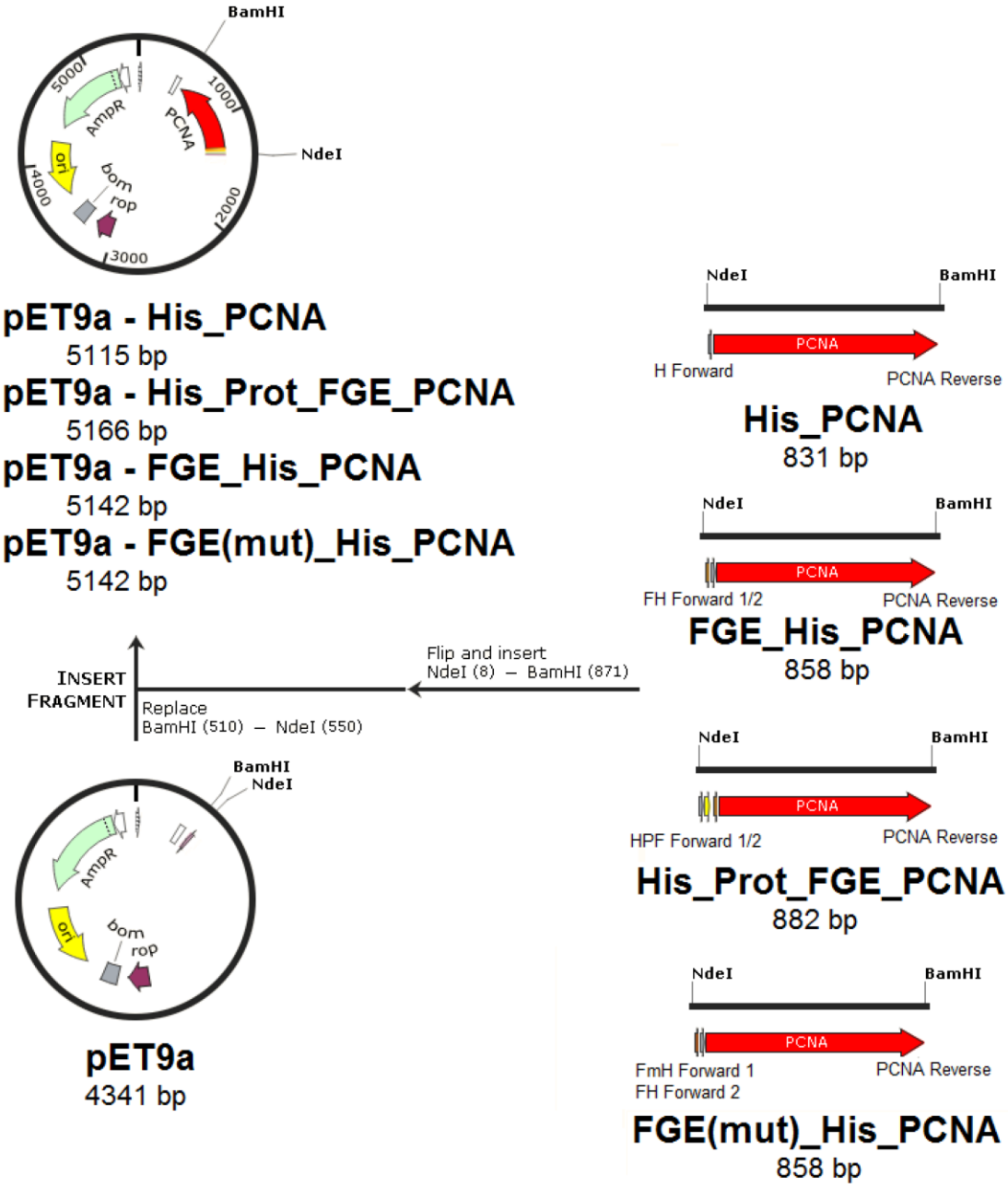

Figure 7. pET9a PCNA Constructs. PCNA was cloned into pET9a (Kan^r) at BamH1 and Nde1 sites. My four constructs are shown on the right. The final constructs are shown at top. Constructs were verified via sequencing.

PCNA Amplification Products

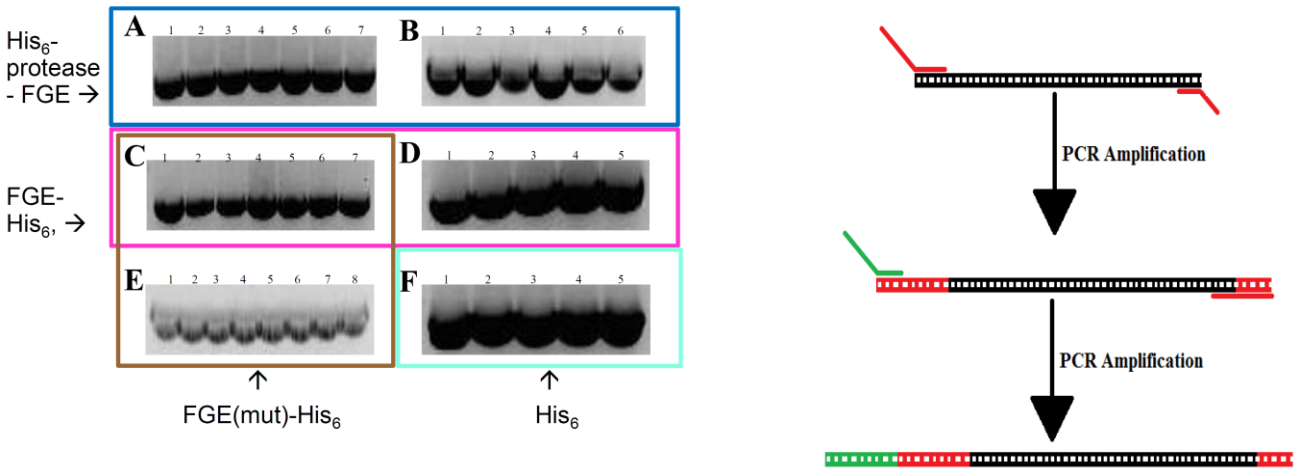

Figure 8. PCNA PCR Products. Agarose gels showing amplification of PCNA after the addition of partial tags. On the right is the dual-primer strategy used to attach longer tags. The gene is initially amplified using a set of inner primers. The product of this reaction is then amplified with an outer set of primers to generate the complete tags. **A)** Inner primer amplification of His_6, protease, FGE construct. Reaction was performed in bulk and split into 7 lanes. All 7 lanes were excised and purified together. **B)** Using the purified product from A, the outer primers were applied and the complete His_6-protease-FGE tags were finalized. **C)** Results from the amplification of the inner tags of the FGE-His_6 construct. Reaction was performed in bulk and split into 7 lanes. All 7 lanes were excised and purified as one product. **D)** Amplification of product from C with the outer tags to generate the complete FGE-His_6 construct. All lanes were excised and purified prior to digestion and subcloning. **E)** FGE(mut)-His_6 outer tags applied to the product from C. **F)** Single amplification necessary to generate His_6 tag construct. The reaction was done in bulk and split into 5 lanes. All 5 lanes were excised and purified together prior to digestion and cloning into vectors.

RecJ Fusion Tag Design

Mutagenesis Primers		
+6 – xba1-RBS-8bases-gene (w/o stop) -gly-his-gly-k-gly- fge-stop-bamHI-+6	C-terminus His_6 and FGE	
Forward	gcgcgc- tctaga – aaggag – atatacat-ATGAAACAACAGATACAACTTCGTC	Tm= 61 C
Reverse 1	TCATCGACAATATCTGGCCAATT- ggc – cat cat cat cat cat cat – ggc – aaa – ggc	Tm= 61 C
Reverse 2	TCATCATCATCATGGCAAAGGC-ctg tgc acc ccg agc cgt – taa – ggatcc - gcgcgc	Tm= 61 C
Sequencing Primers		
RecJ Seq 1-JB	cgg cag act tgc ccg	
RecJ Seq 2-JB	aac atc gca att cct	
RecJ Seq CTerm-JB	gct gaa gtt tgg cgg	
RecJ Seq Nterm-JB (RC)	cgc gaa gga acg cgg	

Figure 9. RecJ Primer Constructs. RecJ fusion constructs have C-terminal His_6 and FGE tags. The length of the tags required the use of 2 primers for efficient PCR amplification. They are labeled reverse 1 and 2; reverse 1 being used first and reverse 2 used on the product from amplification with reverse 1. The same forward primer was used both times. Beneath that are sequencing primers. The sequencing primers were used to verify the construct after cloning into pET9a/3a.

RecJ 63.389 kD Primer allocation

gtgaaacaacagatacaacttcgtcgccgtgaagtcgatgaaacggcagacttgcccgctgaattgcctcccttgctgcgccgt
ttatacgccagccggggagtacgcagtgcgcaagaactggaacgcagtgttaaaggtatgctgccctggcagcaactgagcg
gcgtcgaaaaggccgttgagatcctttacaacgctttcgcgaaggaacgcggattattgtggtcggtgatttcgacgccgacg
gcgcgaccagcacggctctaagcgtgctggcgatgcgctcgcttggttgcagcaatatcgactacctggtaccaaaccgtttcg
aagacggttacggcttaagcccggaagtggtcgatcaggcccatgcccgtggcgcgcagttaattgtcacggtggataacggt
atttcctcccatgcgggggttgagcacgctcgctcgttgggcatcccggttattgttaccgatcaccatttgccaggcgacacatt
acccgcagcggaagcgatcattaaccctaacttgcgcgactgtaatttcccgtcgaaatcactggcaggcgtgggtgtggcgtt
ttatctgatgctggcgctgcgcaccttttgcgcgatcagggctggtttgatgagcgtaacatcgcaattcctaacctggcagaac
tgctggatctggtcgcgctggggacagtggcggacgtcgtgccgctggacgctaataatcgcattctgacctggcaggggatg
agtcgcatccgagccggaaagtgccgtccggggattaaagcgctgcttgaagtggcaaaccgtgatgcacaaaaactcgccg
ccagcgatttaggttttgcgctggggccacgtctcaatgctgccggacgactggacgatatgtccgtcggtgtggcgctgttgtt
gtgcgacaacatcggcgaagcgcgcgtgctggcaaatgaactcgatgcgctaaaccagacgcgaaaagagatcgaacaag
gaatgcaaattgaagccctgaccctgtgcgagaaactggagcgcagccgtgacacgctacccggcgggctggcaatgtatca
ccccgaatggcatcagggcgttgtcggtattctggcttcgcgcatcaaagagcgttttcaccgtccggttatcgcgtttgcgcca
gcaggtgacggtacgctgaaaggttccggtcgctccattcaggggctgcatatgcgtgatgcgctggagcgattagacacact
ctaccctggcatgatgctgaagtttggcggtcatgcgatggcggcgggtttgtcgctggaagaggataaattcaaactctttcaa
caacggtttggcgaactggttactgagtggctggacccttcgctattgcaaggcgaagtggtatcagacggtccgttaagcccg
gccgaaatgaccatggaagtggcgcagctgctgcgcgatgctggcccgtgggggcagatgttcccggagccgctgtttgacg
gtcatttccgtctgctgcaacagcggctggtgggcgaacgtcatttgaaggtgatggtcgaaccggtcggcggcggtccactg
ctggatggtattgcttttaatgtcgataccgccctctggccggataacggcgtgcgcgaagtgcaactggcttataagctcgatat
caacgagtttcgcggcaaccgcagcctgcaaattatcatcgacaatatctggccaatttag

Figure 10. RecJ Primer Allocation. Placement of sequencing primers from figure 9 in RecJ. This assured complete gene coverage for sequencing.

RecJ Constructs

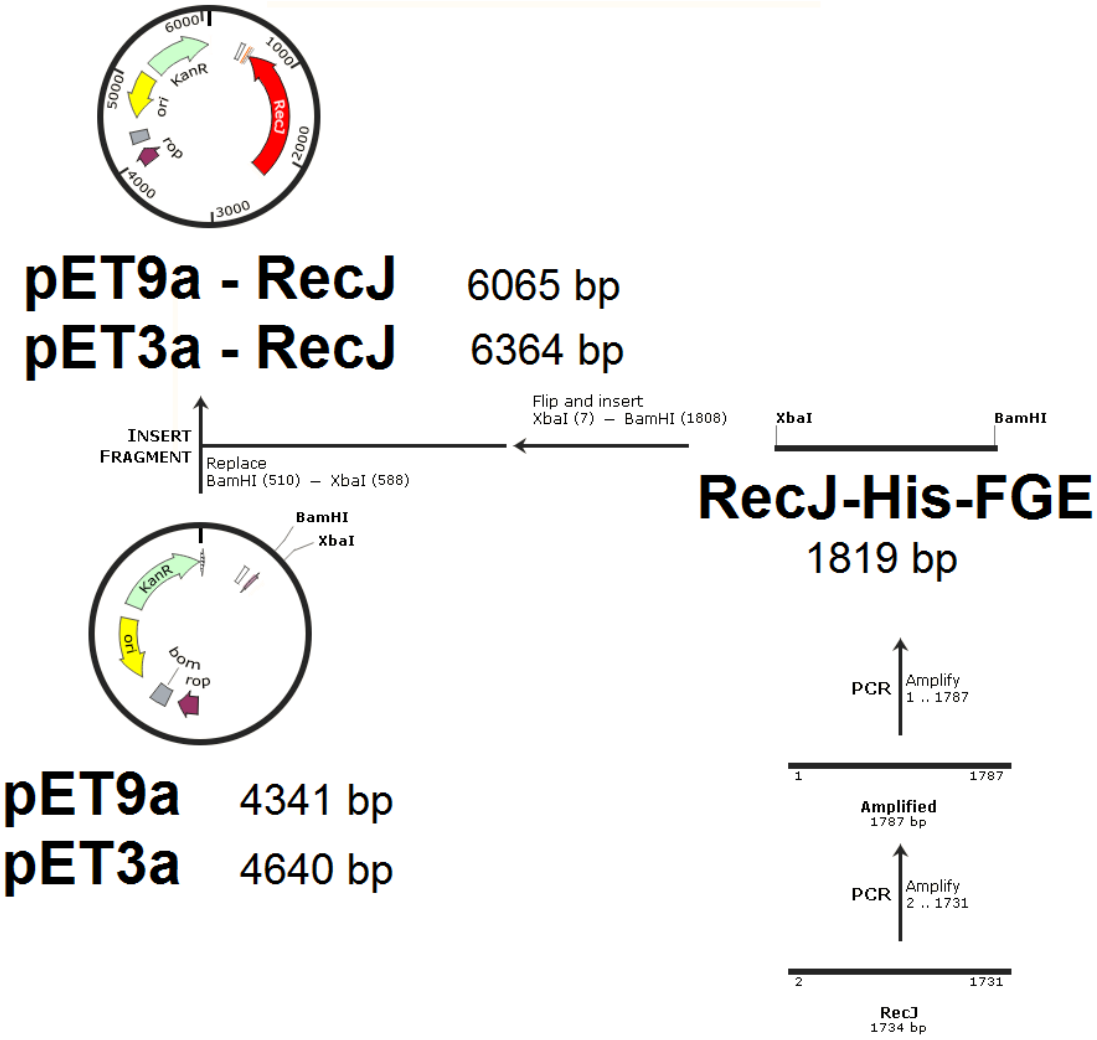

Figure 11. RecJ Construct Design. The final constructs for RecJ were generated by insertion of amplified RecJ into the BamH1 and Xba1 sites of pET9a and pET3a. The two amplification steps are clearly visible on the right, corresponding to the two primers needed to attach the full-length tags. This final construct was digested with BamH1 and Xba1 to be compatible with the vectors. After ligation, constructs were verified by sequencing using the primers shown in figure 9.

RecJ Amplification Products

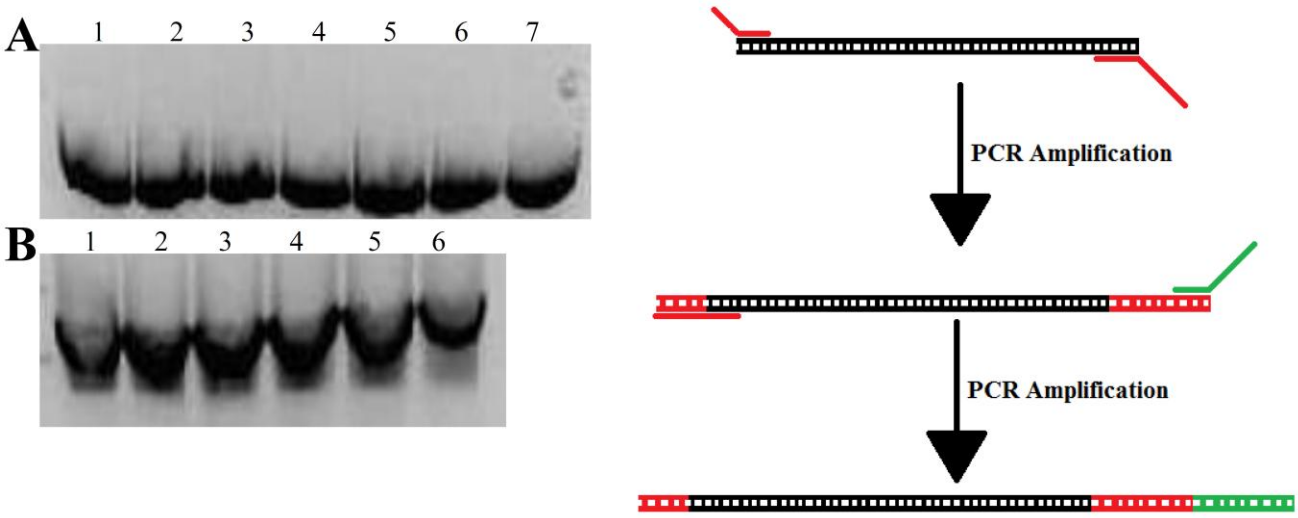

Figure 12. RecJ PCR Products. Agarose gel electrophoresis of PCR products from RecJ reactions. The dual-primer design is outlined at right. The wild type gene was amplified using an inner set of primers. The product of this reaction was then amplified using an outer set of primers to complete the fusion tags. **A)** RecJ amplified with the inner tags. Bulk reaction was split into 7 lines prior to electrophoresis. All 7 lanes were excised and purified together. **B)** Final RecJ insert, generated from the outer tags applied to the product from A. These 6 lanes were purified together prior to digestion and ligation into vectors.

FGE Tag

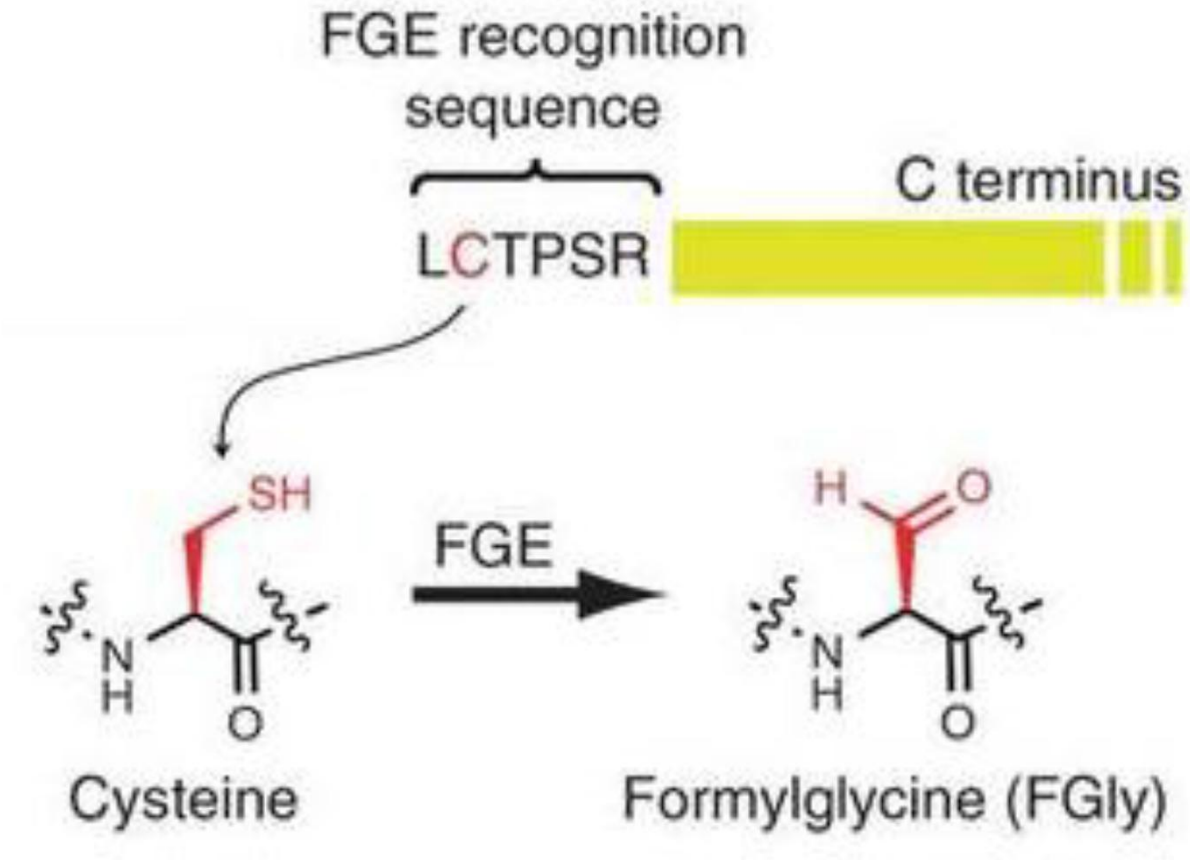

Figure 13. FGE Tag Recognition and Reaction. FGE recognizes a hexa-amino tag, LCTPSR, and converts the central cysteine to a formylglycine. In the FGE(mut) PNCA, this cysteine is replaced with an alanine. This allowed us to test the specificity of FGE conversion.

Chapter 2

Purification and Fluorophore Labeling of Select Mismatch Repair Proteins

Introduction:

For MutS to be used in single-molecule Total Internal Reflection Fluorescence (TIRF) studies, it had to be over expressed and purified. There are several methods to express engineered proteins. Bacterial expression provides a fast and efficient method for expression (Miroux and Walker, 1996). However, bacterial cells lack post-translational modifications common in eukaryotic cells, and some eukaryotic proteins are toxic in bacteria (Terpe, 2006). Yeast cultures grow relatively fast and also provide large cultures to express proteins. Howeverinsect cells provide an incomplete assortment of post-translational modifications to human proteins(Cregg et al., 2000). Insect cells provide a third method for protein expression. Often, recombinant proteins are soluble and non-toxic to the cells. Insect cells grow much slower than bacterial or yeast cells and require transfection and baculovvirus isolation (Kost et al., 2005). For my purposes, MutS and RecJ could be expressed in bacterial cells.

PCNA expression was examined in bacterial cells. Ideally, PCNA would be expressed in human cells to maintain the appropriate post-translational modifications. However, human tissue cultures are extremely difficult and expensive to work with (Alberts et al., 2002). Moreover, PCNA appears to be modified throughout the cell cycle with a variety of post-translational modifications that may confuse single molecule analysis, interfering with desired protein-protein interactions. Clearly, the effect of such post-translational modifications on MMR and other PCNA-dependent processes will be necessary for future studies. However, unmodified PCNA provides a convenient starting point for these studies.

Protein purification is accomplished through protein chromatography. Traditional purification techniques have involved gel filtration, or size-exclusion, chromatography. This technique passing the proteins through a matrix that separates them based upon size (skoog, 2006). Another method for

separating proteins is affinity chromatography. This method provides a column material on which the protein of interest can be specifically bound and separated from proteins that do not recognize the matrix (Uhlen, 2008). This technique was improved by the creation of fusion tags such as Glutathione-S-transferase (GST), Maltose Binding Protein (MBP), and the His_6 tags. GST, binding to GSH, and MBP, binding to amylose, may act as affinity tags and solubility tags (Nallamsetty and Waugh, 2007). However, both these tags are large. His_6 tags are small (six amino acids) that coordinate 2+ metal ion column, such as Ni^{2+}. Proteins bound to Ni^{2+} in a column matrix may be eluted by the addition of imidazole, a histidine analog (Hengen, 1995). This makes it nearly ideal for purification and is unlikely to affect protein diffusion in the single-molecule experiments.

Methods:

Protein overexpression – The MutS, PCNA, and RecJ overexpression plasmids were transformed into BL21(DE3) pLysS bacterial cells for expression (Merk Millipore, Darmstadt, Germany). The DE3 designation means the strains contain the λ DE3 lysogen which carries the gene for T7 RNA polymerase under control of the lacUV5 promoter. The pLysS plasmid produces T7 lysozyme to reduce basal level expression of the gene of interest. The T7 promoter is under the control of Isopropyl β-D-1-thiogalactopyranoside (IPTG) and regulation by LacA/LacO. Addition of IPTG causes LacA to dissociate from the LacO promoter, allowing for expression of a gene of interest. Single colony transformants were grown in 1X LB, shaking at 240 rpm (Miller's Recipe, EMD Millipore).

MutS - MutS expression was induced with 1μM IPTG after growing to an A_{600} = 0.1 and further grown overnight shaking at 17°C. Cultures were precipitated by centrifugation at 4,000 rpm for 15 min (Sorval Legend XTR centrifuge, rotor #75003607) at 4°C. The cell pellet was resuspended in 25mM Hepes (pH 7.5), 300mM NaCl, 10% Glycerol, and 0.384ug/ml Pepstatin, 0.384ug/ml Leupeptin, and 0.24 mM PMSF (protease inhibitors) (Sigma-Aldrich, St Louis, MO) and pelleted by centrifugation as above. The second cell pellet was resuspended in25mM Hepes (pH 7.5), 300mM NaCl, 10% Glycerol, 0.384ug/ml Pepstatin, 0.384ug/ml Leupeptin, 0.24mM PMSF (Sigma-Aldrich, St Louis, MO) and 200mM Imidazole (Sigma-Aldrich, St Louis, MO), frozen and stored until use at -80°C.

10 mL fractions for solubility tests were spun down at 4,000 rpm for 10 min (Sorval Legend XTR centrifuge, rotor #75003607) at 4^{o}C. They were resuspended in 400ul of 1x Dulbecco's Phophate Buffered Saline (DPBS) (Sigma-Aldrich, St. Louis, MO). Cells were lysed by 2 sessions of sonication on ice (40% amplitude for 30 sec), and spun again at 14,000 rpm for 10 min. Supernatent was removed and kept as the soluble fraction. The pellet was resuspended in 400ul 400ul of 1x DPBS and centrifuged 2 more times. Final resuspension in 400ul of 1x DPBS and run on a 10% poly-acrilamide gel with the soluble fraction.

Pellets lysed by 3 sessions of freeze(-80^{o}C)-thaw on ice (4^{o}C) and 2 sessions of sonication on ice (30% amplitude for 1 min). Lysis extracts were centrifuged in a Beckman 60 Ti rotor at 4^{o}C and 40,000 rpm for 1 hour. The supernate (Fraction I) was applied to a 1 ml Ni-NTA column (GE Healthcare) at 0.15 ml/min and eluted with a 200mM Imidazole step (Fraction II). Purified FGE protein was added to Fraction II containing MutS (3:1 molar ratio) and the mixture dialyzed against conversion buffer (50 mM Tris pH 8, 100mM NaCl, 20mM arginine, 0.25mM DTT) for 48 hrs (Fraction III). Fraction III was dialyzed against labeling buffer (100 mM Potassium Phosphate pH6.5, 300 mM NaCl, 0.25 mM DTT) overnight (Fraction IV). HIPS-Alexa647 (0.5 mM in labeling buffer; Redwood Biosciences, Emeryville, CA)) was added to Fraction IV and labeling continued for 48 hr at 0^{o}C (ice water bath). Following labeling (Fraction V), the protein fraction was diluted 3x with 25 mM Hepes pH 7.5, 1 mM DTT, 0.1 mM EDTA, 10% glycerol and applied to a 1 ml Heparin column (GE Healthcare) at 0.2 ml/min. The column was washed with 25 mM Hepes pH 7.5, 100 mM NaCl, 1 mM DTT, 0.1. mM EDTA, 10% glycerol and eluted with a 30 ml gradient to 1 M NaCl in the same buffer. MutS eluted at approximately 0.35 M NaCl (Fraction VI; ~12 μM peak fractions). Fraction VI was dialyzed against 25 mM Hepes pH 7.5, 150 mM NaCl, 1 mM DTT, 0.1 mM EDTA, 20% glycerol, aliquoted (~10 μl), flash frozen in liquid N_2 and stored at -80^{o}C.

PCNA - PCNA was induced with 0.1 mM IPTG after growing to an A_{600} = 0.3 and further grown overnight shaking at 37^{o}C for 8 hours. Cultures were centrifuged at 3,500 rpm for 15 min (Sorval Legend XTR centrifuge, rotor #75003607) at 4^{o}C. The cell pellet was resuspended in 25 mM Hepes (pH 7.5), 150 mM NaCl, 10% Glycerol, and 0.384ug/ml Pepstatin, 0.384ug/ml Leupeptin, and 0.24 mM PMSF (Sigma-Aldrich, St Louis, MO) and pelleted by centrifugation as above. The second cell pellet was resuspended in25mM Hepes (pH 7.5), 300mM NaCl, 10% Glycerol, 0.384ug/ml Pepstatin,

0.384ug/ml Leupeptin, 0.24mM PMSF (Sigma-Aldrich, St Louis, MO) and 200 mM Imidazole (Sigma-Aldrich, St Louis, MO), frozen and stored until use at -80^{o}C.

RecJ - RecJ was induced at 0.01mM IPTG after growing to an A_{600} = 0.5 and further grown overnight shaking at 16^{o}C. Cultures were centrifuged at 3,500 rpm for 15 min (Sorval Legend XTR centrifuge, rotor #75003607) at 4^{o}C. The cell pellet was resuspended in 25 mM Hepes (pH 7.5), 150 mM NaCl, 10% Glycerol, and 0.384ug/ml Pepstatin, 0.384ug/ml Leupeptin, and 0.24 mM PMSF (Sigma-Aldrich, St Louis, MO) and pelleted by centrifugation as above. The second cell pellet was resuspended in25mM Hepes (pH 7.5), 300mM NaCl, 10% Glycerol, 0.384ug/ml Pepstatin, 0.384ug/ml Leupeptin, 0.24mM PMSF (Sigma-Aldrich, St Louis, MO) and 200 mM Imidazole (Sigma-Aldrich, St Louis, MO), frozen and stored until use at -80^{o}C.

Pellets lysed by 3 sessions of freeze(-80^{o}C)-thaw on ice (4^{o}C) and 2 sessions of sonication on ice (30% amplitude for 1 min). Lysis extracts were centrifuged in a Beckman 60 Ti rotor at 4^{o}C and 40,000 rpm for 1 hour. The supernate (Fraction I) was applied to a 1 ml Ni-NTA column (GE Healthcare) at 0.15 ml/min and eluted with a 200 mM Imidazole step (Fraction II). Purified FGE protein was added to Fraction II containing MutS (3:1 molar ratio) and the mixture dialyzed against conversion buffer (50 mM Tris pH 8, 100 mM NaCl, 20 mM arginine, 0.25 mM DTT) for 48 hrs (Fraction III). Fraction III was dialyzed against labeling buffer (100 mM Potassium Phosphate pH6.5, 300 mM NaCl, 0.25 mM DTT) overnight (Fraction IV). HIPS-Atto488 (0.5 mM in labeling buffer; Redwood Biosciences, Emeryville, CA) was added to Fraction IV and labeling continued for 48 hr at 0^{o}C (ice water bath). Labeling efficiency was examined by PAGE and compared to intrinsic labeling of the FGE protein.

Complementation of the RecJ fusion protein constructs were performed in *E.coli* AB1157 Δ*RecJ* on LB agar. Complementation was tested using U.V. exposure (Biorad transilluminator) of streaked cells on an LB plate for 3, 6, 9, 12, and 15 sec (Wang and Smith, 1988).

Results:

MutS expression, purification, and labeling - MutS was exceptionally difficult to express as a soluble protein. At generally standard growth conditions, 37^{o}C and 0.1-0.5 mM IPTG, MutS was mostly

insoluble (Figure 1). It required several trials of IPTG concentrations and growth temperature before reasonable expression condition was found: initial growth to A_{600} = 0.1, addition of 1 μM IPTG, and further growth overnight at 17°C (Figure 2). These conditions allowed for the most soluble MutS by slowing down the expression. However, with slower expression (lower IPTG or temperature), longer times were needed to produce enough protein in culture to purify. Thus, expressions had to be run for 12-16 hrs (overnight).

Purification and labeling of MutS was largely successful, although contaminating proteins that appear less than 3% of the total protein as can be seen in both the coomassie stained gel and the typhoon (Molecular Dynamics) visualization of labeled proteins (Figure 3). MutS was co-expressed with FGE to provide *in vivo* conversion of the FGE tag since we found that co-expression of the FGE protein with MutS in the *E.coli* cells resulted in less than 30% conversion of the central cysteine in the FGE tag (data not shown). Following *in vitro* FGE conversion, incubation with HIPS-Alexa647 and separation of the labeled protein from the unincorporated dye using Heparin column chromatography, we observed that 26% of the total MutS monomers were labeled with Alexa647 (Figure 3). Because MutS functions as a dimer, this amount of labeling translates to 45% of the functional protein. Additional work in the laboratory has increased the dimer labeling to as high as 75%. However, the probability of proteins that contain *two* fluorophore labels increases multiplicatively as well. Studies in the laboratory have suggested that a high background of multiply labeled MutS may confuse the observation of multiple sliding clamps during single molecule analysis. Thus, the slightly lowered labeling efficiency for the bacterial MutS protein is advisable.

PCNA expression - PCNA was significantly simpler to express than MutS (Figure 4). We found that it expressed well at 37°C with 0.1 mM IPTG added at A_{600} = 0.3 and then an additional growth at 37°C shaking for 8 hrs (Figure 4A). Moreover, the PCNA protein was largely soluble under these overexpression conditions (Figure 4B). We have not as yet performed successful protein purification and fluorophore labeling. Initial studies used the His_6-FGE-PCNA construct. Unpublished observations in the laboratory have suggested that an external FGE tag with an internal His_6 tag may be more amenable to labeling. These results were the foundation for the construction of the His_6-HRV-FGE-PCNA and FGE-His_6-PCNA, where the external His_6-tag may be removed by the HRV-3C protease or the FGE is external to the His_6-tag. The FGE(mut)-His_6-PCNA was constructed to

determine the specificity of labeling since it is clear from Figure 3 that protein *not* containing an FGE-tag appear to be labeled by the HIPS-Alexa647. Dr. Jiaquan Liu has performed similar specificity studies using an FGE(mut)-His_6-MutS construct and purified protein and found it to be >95% specific for the wild type FGE-tag (JL personal communication).

RecJ complementation **-** One of the advantages to examining bacterial genes/proteins is that one can determine whether altered gene fusion constructs are able to complement bacteria containing deletions of the bacterial gene through easily determined phenotypes (Figure 5). Because *recJ* deletion strains are sensitive to UV, a simple qualitative UV exposure assay was developed (Figure 5). In this system, overnight liquid cultures that represent tests and controls are diluted to identical concentration of cells as determined by A_{600}. These cultures are then streaked on an LB plate using a cotton swab and exposed to UV for various times (Figure 5). As expected *ΔrecJ* cells are sensitive to UV Light compared to the isogenic *wild type* parent bacterial cell (AB1157). Complementation analysis demonstrated that the FGE-His_6-RecJ fusion construct appears UV resistant when transformed into *ΔrecJ* cells. These results suggest that the FGE-His_6-RecJ fusion construct retains function, even with the fusion tags. As expected, *wild type* cells are largely unaffected by UV light. Interestingly, *wild type* cells containing the FGE-His_6-RecJ fusion construct appeared UV sensitive. This would appear to imply that overexpression of RecJ is not beneficial, implying a dosage effect. Studies of RecJ expression levels in a more quantitative UV cell survival assay will be required to fully detail this phenomenon.

RecJ expression **-** RecJ expressions were complicated by the insolubility of the protein (Figure 6). It is insoluble at standard expression conditions (37^oC, 0.1-0.5 IPTG). However, building on my experience with MutS, similar expression conditions were found. Final conditions for making pellets are: 17^oC, O.D. 0.5, 1 μM IPTG (Figure 6). Expressions have to be run 12 - 16 hrs to produce enough protein for purification, and even then two-liter pellets are required to produce enough soluble RecJ. In conjunction with expression tests, complementation assays were performed.

RecJ labeling **-** Once *in vivo* complementation was shown, and expression conditions determined, RecJ could be purified and labeled. A test purification using only a nickel column was performed (Figure 7). Purification was successful with only minor contaminants remaining. Fractions 18 to 32 were then

used to test labeling. *In vitro* conversion with FGE was done and then RecJ was labeled by Cy3 from Redwood biosciences (Figure 8). In figure 8, the left lanes show RecJ with no FGE. This RecJ should undergo little to no conversion. The right lane had FGE added, to perform conversion, and shows superior labeling to the unconverted lane. However, labeling was less than 10% efficient, not enough for use in single-molecule studies. Labeling efficiency was quantified using FGE as the standard. FGE is auto-catalytic and becomes labeled during testing. Comparing the amount of labeled FGE to labeled RecJ, knowing the amount of each protein in the reaction, we could determine approximately how much RecJ was labeled. Another member of the lab switched the tags to the N-terminus of the protein and repeated labeling. This time, RecJ was co-expressed with FGE to allow for *in vivo* conversion of the FGE tag (Figure 9). Figure 9 shows RecJ labeling with and without added FGE. Since there is no difference between the lanes, as much of the protein was labeled after *in vivo* conversion as with some *in vitro* conversion. Thus, it had complete conversion *in vivo*. This makes purification and labeling simpler later.

Discussion:

MutS **-** MutS provided a test case for generating and labeling FGE tagged proteins. Successful expression and purification allowed for labeling tests to be conducted. Members of the Fishel laboratory have used the labeled protein in single-molecule studies. In these studies, diffusion of the MutS during the mismatch search as well as following the formation of an ATP-bound sliding clamp was examined. These studies prove the efficacy of our method for labeling proteins and its use in single-molecule studies.

PCNA **-** The process of FGE tagging and labeling procedures was extended to PCNA. PCNA expression conditions were less stringent than MutS expressions. This allowed for cultures to be grown at a higher temperature and expressions lasted 7 hours. Expression experiments were conducted on His-PCNA. A future direction will be to determine expression conditions for other PCNA constructs followed by purification of FGE-tagged PCNA for fluorophore labeling experiments.

RecJ **-** RecJ expression was complicated by the low temperature and concentration of IPTG necessary to produce soluble protein. Final conditions of 17^{o}C and 5 μM IPTG were determined to produce the largest fraction of soluble protein. However, this necessitated over-night expressions and purification

of at least two-liters of cells at a time. Purification was straightforward and successful. RecJ was purified with small amounts of contamination over only a nickel column. In the future, heparin (nucleotide analog) or ion exchange chromatography will be used in conjunction with the nickel column. Initial labeling tests of C-terminal labeled RecJ showed it to be a poor substrate for FGE conversion or labeling. Tests showed less than 10% labeling of RecJ. Rearranging the tags to the N-terminus and co-expressing with FGE, for *in vivo* conversion, allowed for complete conversion of the tag. Subsequent experiments will determine the efficiency of the labeling. The next step is to test the activity of RecJ and then use it in single-molecule studies.

MutS Expression

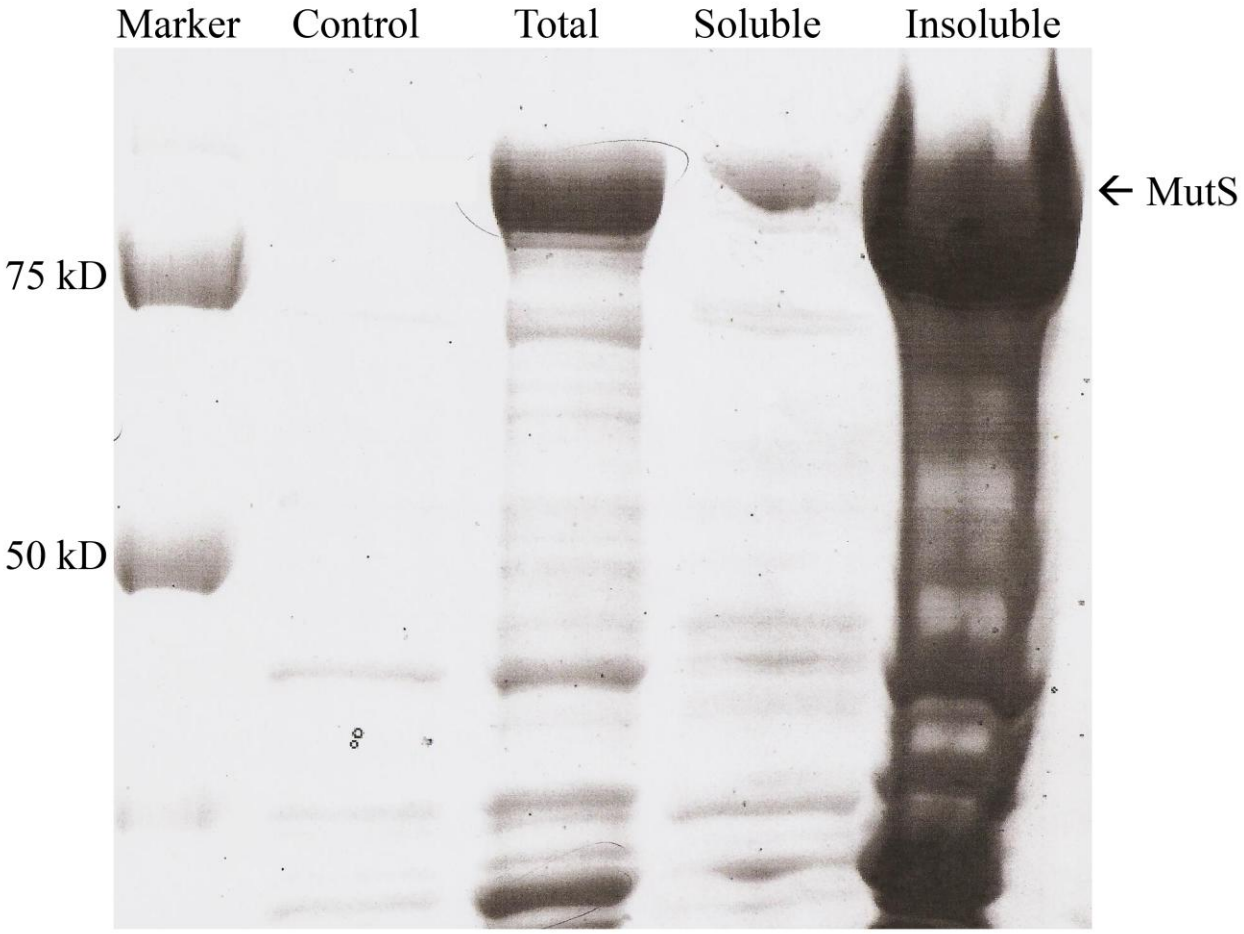

Figure 1. Early MutS Expression. Initial expressions of MutS, at 37^{O}C and induced with 0.1 mM to 0.5mM IPTG, produced large amounts of mostly insoluble protein. The middle lane shows total cell lysate, followed by the soluble and insoluble fractions. While there is some soluble MutS, the vast majority is insoluble.

MutS and FGE Co-Expression

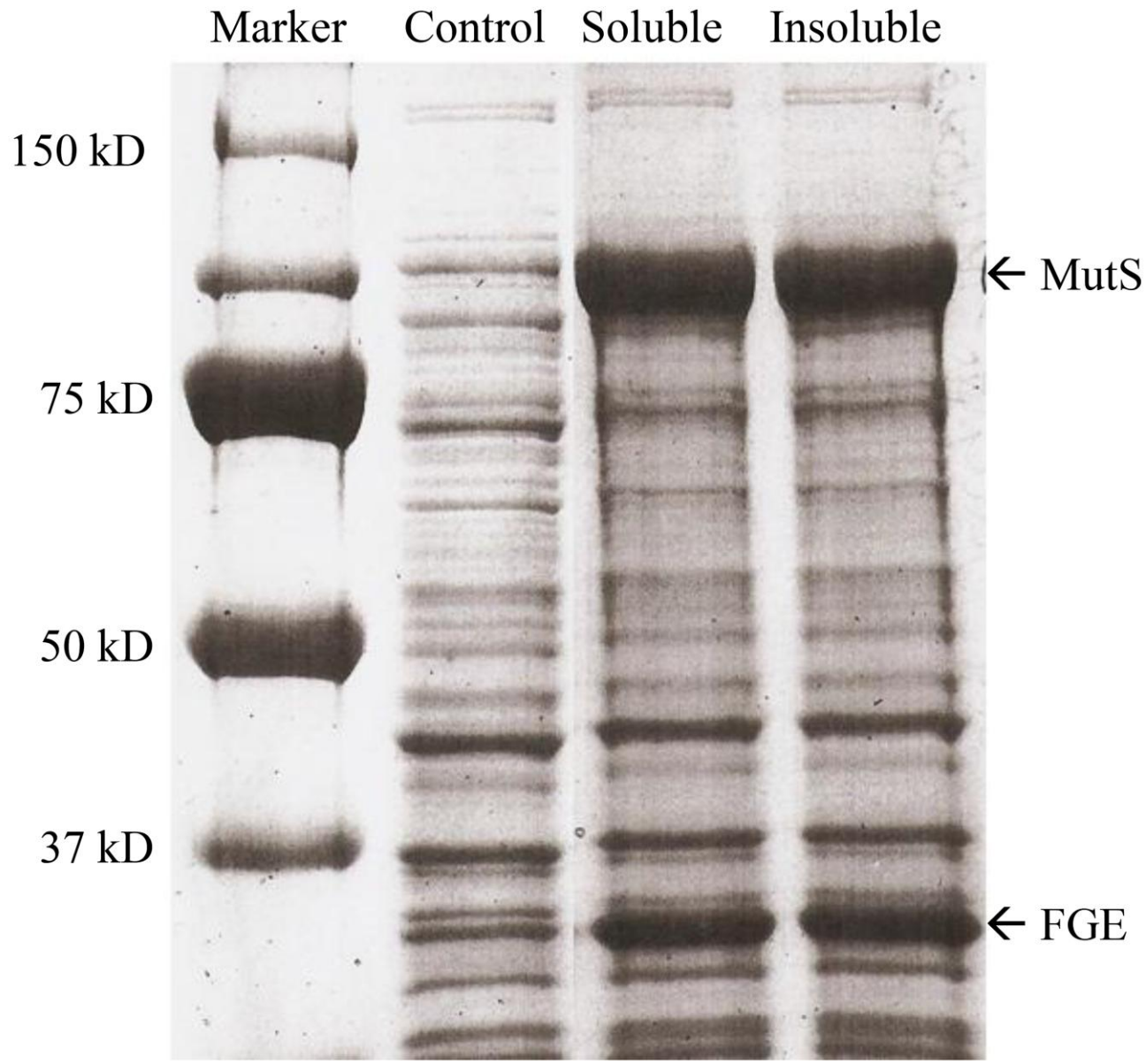

Figure 2. MutS and FGE Co-expressions. This expression was performed at 17^{O}C, 1 µM IPTG overnight. The control lane is a sample from right before induction. The soluble and insoluble fractions were taken after overnight expression. There is more soluble MutS here than in previous expressions and this was chosen as the final conditions for making pellets.

MutS Expression and Labeling

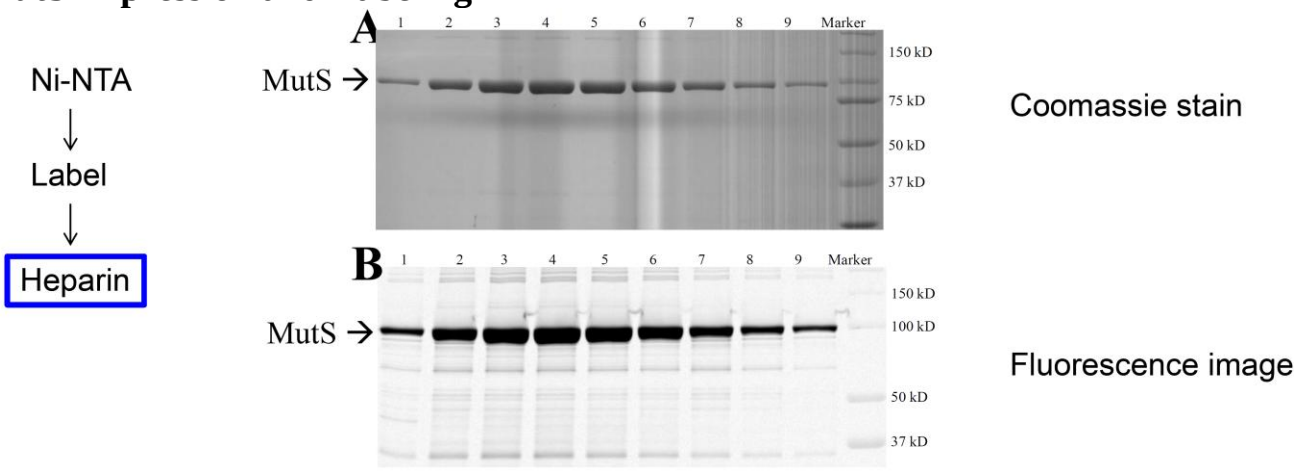

Figure 3. MutS Expression and Labeling. Co-expression of MutS with FGE allowed for *in vivo* conversion of the FGE tag, eliminating the need for conversion after purification. Thus, MutS was labeled during purification, prior to being run over the final column, a Heparin column. **A)** A coomassie stain of labeled MutS. Fractions 1 through 9 are peak fractions of MutS from the Heparin column. **B)** Fluorescent picture of figure A, the wavelength of emission is 665 nm. MutS was labeled with Alexa 647.

PCNA Expression

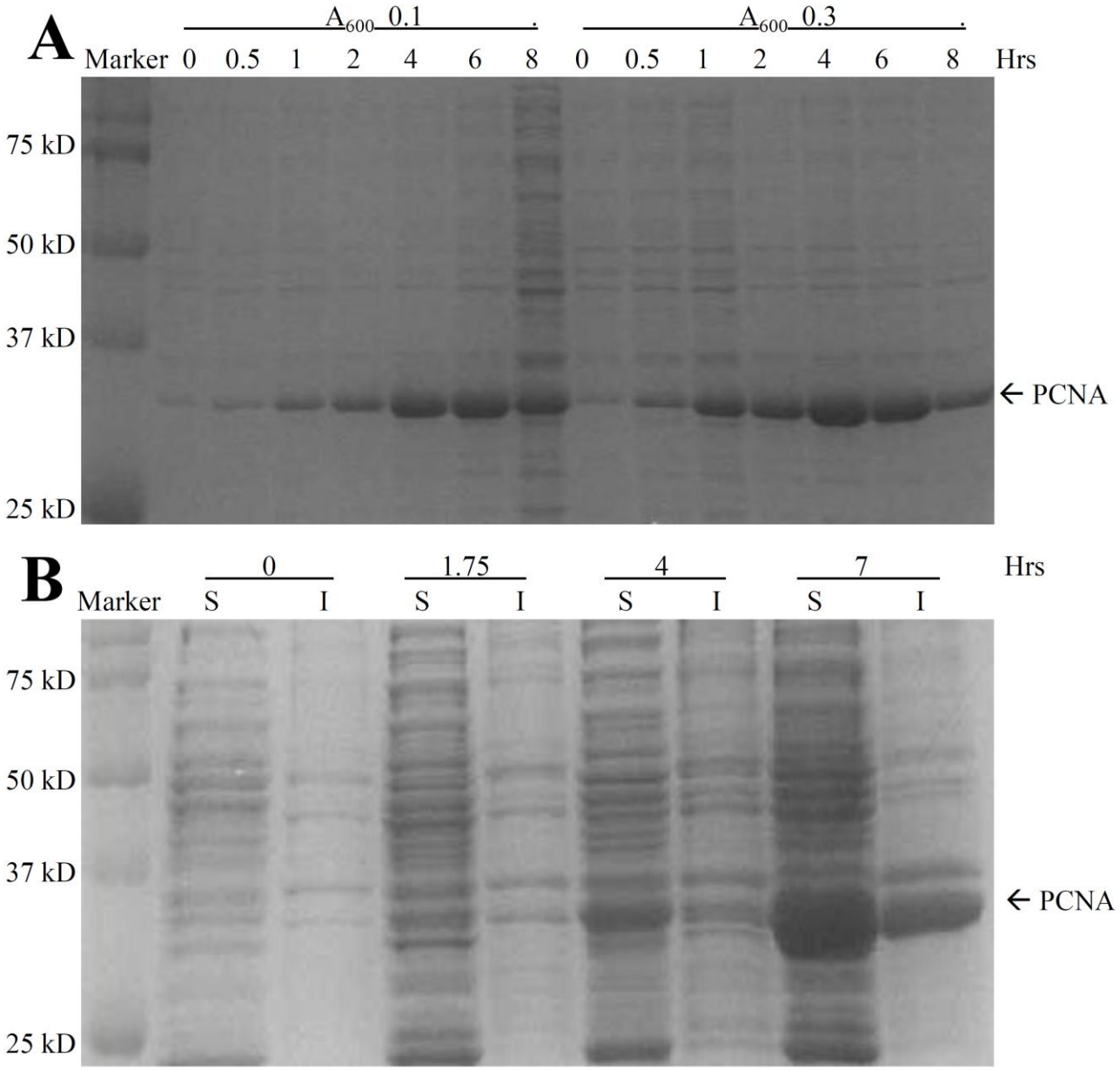

Figure 4. PCNA Expressions. PCNA expressions were successful under less stringent expression conditions than MutS. **A)** Initial expression performed at 37^{O}C and 0.1mM IPTG. Time points (labeled in hours after induction) for two cultures are shown: A_{600} 0.1 and A_{600} 0.3. **B)** PCNA expressed under the final conditions, 0.4 A_{600}, 0.01mM IPTG, and 37^{O}C, showing soluble and insoluble portions at different time points. This generated enough soluble protein for purification.

RecJ Complementation Assay

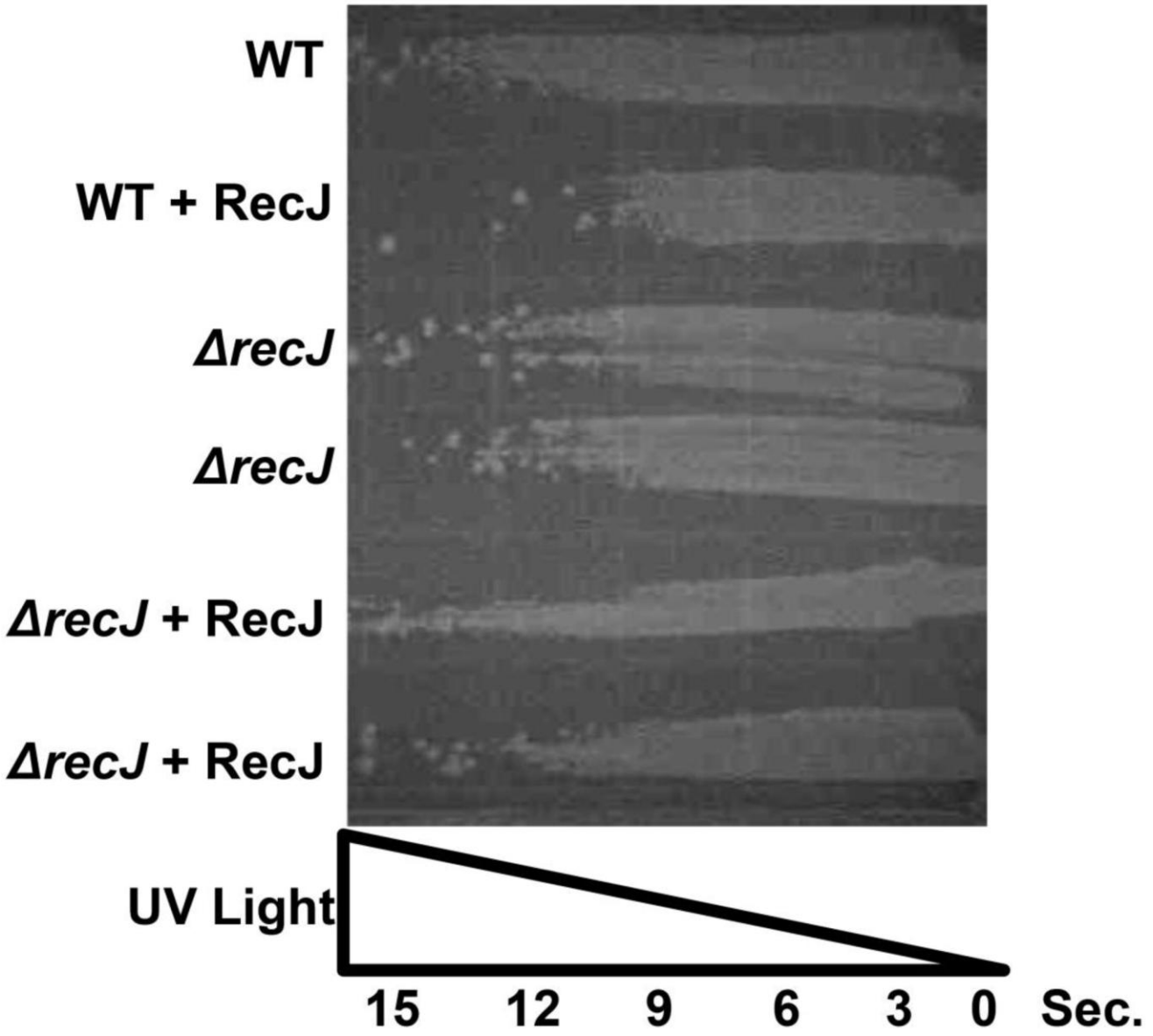

Figure 5. RecJ Complementation. Complementation assays were performed with RecJ using AB1157 *ΔrecJ* cells. RecJ deficient cells are more susceptible to U. V. damage than wild-type strains. Wild-type cells are resistant to U.V. damage. Deficient cells are greatly impacted by U.V. light. Transformed strains show rescue. Interestingly, the wild-type cells with my plasmid are harmed by U.V. light. This would imply a dosage dependence on RecJ.

RecJ Expression

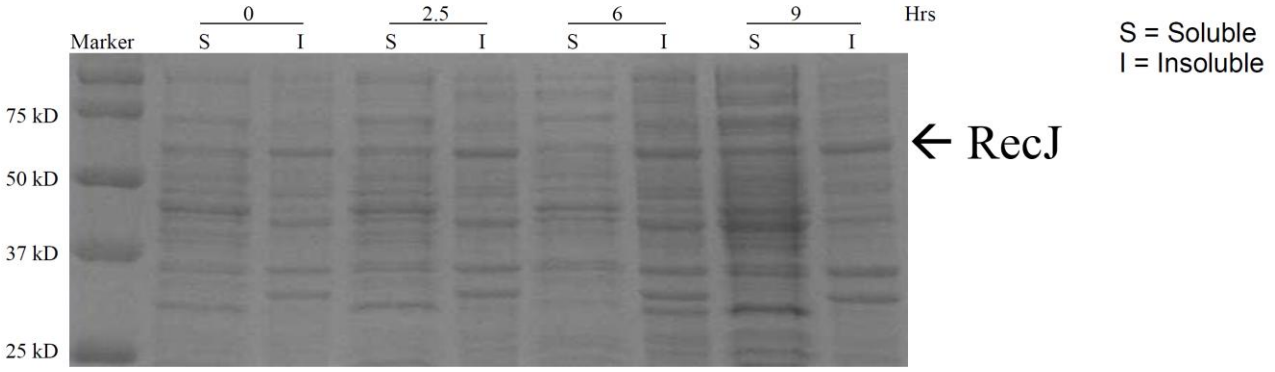

Figure 6. RecJ Expression. RecJ expressions produced mostly insoluble protein at at 37^{o}C and higher (0.1-0.5 mM) concentrations of IPTG. This figure is of soluble and insoluble fractions for different time points at the final expression conditions: 17^{o}C, A_{600} 0.5. and 0.005mM IPTG. Such low expression conditions required overnight expressions and use of two-liter cultures to produce enough protein for purification.

RecJ Purification

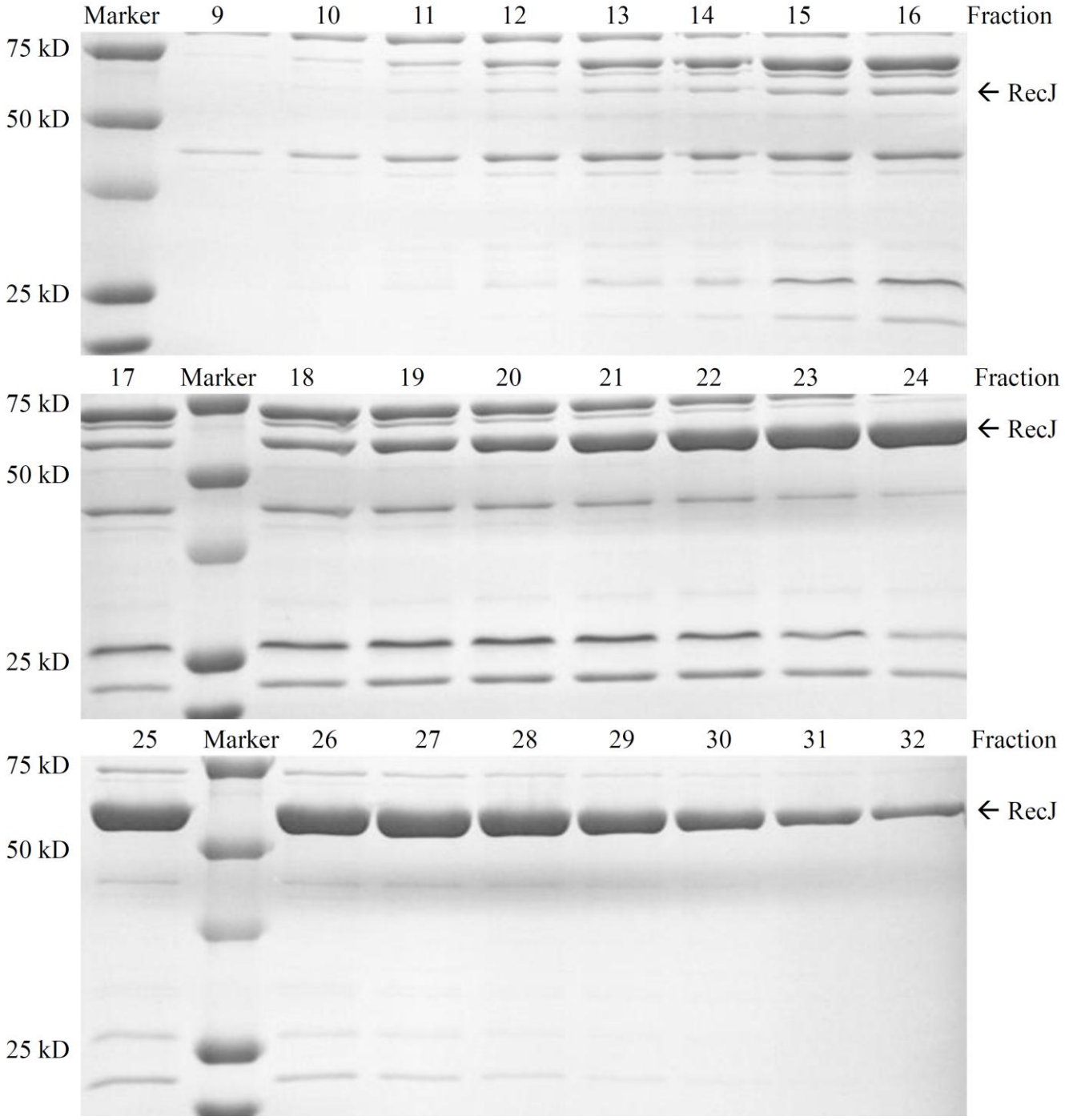

Figure 7. RecJ Purification. The peak fractions after nickel purification of RecJ. RecJ represents the greatest part of the proteins eluted in the peak though with some contamination present. Fractions 18 to 32 were used for labeling tests.

RecJ Labeling: C-Terminal Tags

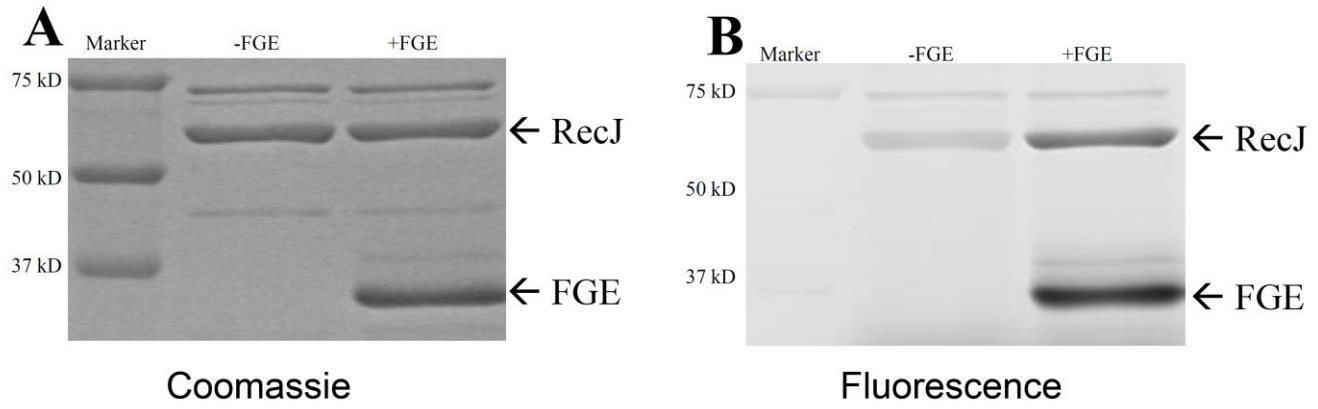

Figure 8. RecJ Labeling Test. *In vitro* conversion of the FGE tag was performed with FGE from a member of the lab. Nickel column fractions were used in this labeling test (see Figure 7). **A)** Coomassie stain of RecJ with and without FGE. Added FGE is clearly visible in the right lane. **B)** A fluorescence picture of A prior to staining. This shows that the specificity of labeling is very high. However, labeling was less than 10% efficient. Cy3 was used for this labeling test. Cy3 has an excitation wavelength of 550 nm and an emission wavelength of 570 nm. This picture was taken using a 570 nm filter.

RecJ Labeling: N-Terminal Tags

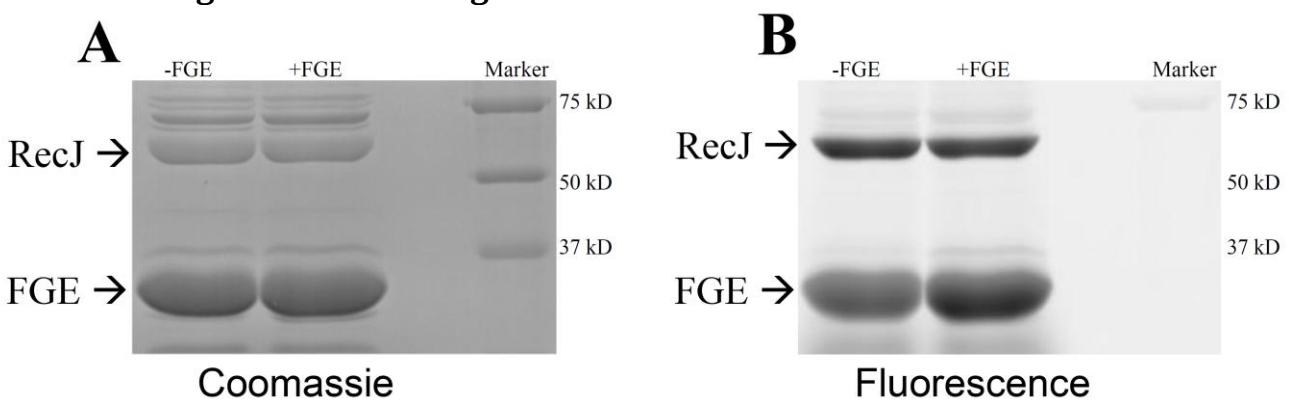

Figure 9. RecJ N-terminal tags. Dr Jiaquan Liu (Fishel Lab) switched the tags from the C-terminus to the N-terminus. FGE was co-expressed with RecJ to allow *in vivo* conversion of the FGE tag. Nickel column fractions were used in this labeling test (see Figure 7) . **A)** Coomassie stain of RecJ, with N-terminal tags, and FGE co-expression. **B)** Fluorescence picture of A shows that labeling is the same, thus complete conversion of the FGE tag *in vivo*. Each right lane had added FGE to test *in vivo* conversion efficiency. Cy3 was used for this labeling test. Cy3 has an excitation wavelength of 550 nm and an emission wavelength of 570 nm. This picture was taken using a 570 nm filter.

Chapter 3

Development of a Radiolabel-free Exonuclease Assay

Introduction:

Purification and labeling exposes proteins to non-native conditions; Temperature changes, pH changes, and ion concentration changes to name a few (skoog, 2006). These can be controlled by purification in a 4^{o}C cold room and carefully maintaining purification conditions similar to cellular conditions. However, this can still have a large effect on an enzyme's activity, so prior to experimentation, activity of the protein must be tested. Traditional enzyme kinetics assays were performed with radio labeled nucleotides (Bilezikian et al., 1975). DNA substrates with P^{32} were generated using αP^{32} labeled dNTPs. These substrates were then digested with the enzyme of interest, and the reactions precipitated with trichloroacetic acid (TCA). TCA precipitates large DNA fragments and proteins associated with them, leaving single nucleotides free in solution. The radioactivity of the solution is then measured and used to determine the amount and rate of digestion. This method is effective but exposes the researcher to harmful radioactivity and contaminates any laboratory materials used in the process. Fluorophores provide an accurate and non-harmful method of visualizing molecules. Using Cy3-dNTPs from GE Healthcare, it is possible to PCR amplify a fragment that is labeled along its entire length (CyDye, Amersham). This allows for a novel enzyme assay that does not rely on radiolabeled materials.

Methods:

DNA generation **-** DNA for enzyme assays was prepared via PCR using Cy3 labeled dCTP (Cy3-dCTP) from GE Healthcare (GE Healthcare, Cleveland, Oh: CyDye protocol, Amersham: Figure 3). Cy3-dCTP was diluted in dCTP for a final ratio of 1:10 Cy3-dCTP:dCTP. This allowed for an average insertion of labeled nucleotides once every ten occurrences of CMP throughout the DNA. Primers for the substrate are shown in Figure 1. DNA was gel purified using 1% agarose gels and visualized using Ethidium Bromide (0.5 ug/ml) under short wave UV fluorescence. Dna was extracted using the QiaGen Gel Extraction kit (Qiagen, Germantown, MD) and quantified using a NanoDrop 2000c

spectrophotometer from Thermo Scientific (Thermo Fisher Scientific, Waltham, MA).

Substrate processing **-** Substrates with a 5' overhang were generated by digestion with endonucleases Apa1 and Avr2, in 1x CutSmart buffer from NEB (Beverly, MA). Reactions were run at 37^{o}C for 90 minutes. Substrates with 3' overhangs were generated by digestion with endonucleases Sac1 and Sph1. Reactions were run at 37^{o}C for 90 minutes in 1x CutSmart buffer from NEB (Beverly, MA).

Nuclease reactions **-** Nuclease reactions were performed in 25mM Hepes (pH 7.8), 1mM DTT, 10mM $MgCl_2$, 1X BSA, and 50mM NaCl at 37^{o}C for 5min, 15min, 30min, 60min, and 120min (Sutera et al., 1999) (Figure 2). Reactions were terminated with 5 ul 50mM EDTA and 10ul of 1mg/ml Salmon Sperm in 1mM EDTA. DNA was precipitated with 15uL 1M Trichloroacetic Acid TCA, resting on ice for 20min, and spun down at 14,000 rpm for 15min in a Microfuge 18 Centrifuge from Beckman Coulter. Reactions were allowed to dark-adapt for 20 min prior to fluorescence measurement by a Fluoromax-4 from Horiba Scientific (Edison, NJ). TCA, aided by Salmon Sperm DNA as a carrier, precipitates large fragments of DNA while leaving single nucleotides in solution. This separates digested substrate from undigested. Using the Fluoromax, the fluorescence of the supernatent, containing digested nucleotides, was measured.

Results:

Kinetics **-** Fluorophores are large additions to DNA nucleotides and thus could have inhibited digestion. Our time-course showed that DNA labeled with fluorophores did not inhibit digestion by RecJ (figure 4). However, we do not know if it affected the rate of digestion. Comparison of our rate with rates determined from traditional, radiolabel, experiments would define the effect of fluorophores on the rate of digestion. We determined that at 0.26 µmol of DNA, RecJ had a linear digestion range of less than 15 min. This knowledge could be extrapolated to design the time course for Michaelis-Menten studies.

Substrate specificity **-** RecJ is a 5'-3' exonuclease (Han et al., 2006). However, nucleases retain function on other substrates, such as single or double stranded DNA or RNA, but are highly preferential towards a single type. This is true of RecJ as well. RecJ preferentially digests ssDNA (figure 5) (Han et al., 2006). After ssDNA, RecJ preferentially digests 5' overhangs on dsDNA, then 3' overhangs of dsDNA, and retains the lowest activity on dsDNA with blunt ends. This is in agreement

with other studies on RecJ (Sutera et al., 1999; Viswanathan and Lovett, 1998).

Michaelis-Menten **-** A Michaelis-Menten study was perfomed to test the kinetics of RecJ. RecJ was incubated with 2, 4, 6, and 8 ng/ul of DNA for 10, 16, 22, and 28 minutes. The results are shown in figure 5, plotted as concentration of DNA vs velocity of reaction. The points were fit to the steady-state Michaelis-Menten equation to determine V_{max}, K_M, and from them, kcat. Vmax, the maximal rate of digestion, was found to be 3.6 µM/min of ssDNA. The Michaelis constant, K_M, is 36.4 µM. This is equivalent to the concentration at half maximal velocity. From this, we determined kcat to be 3.2 s^{-1}. Kcat represents the catalysis, or turnover, rate of the enzyme. The time points used in this experiment, 10 min to 28 min, exceed the linear range determined from the kinetics experiments. These measurements should be repeated, keeping the time to less than 20 min to ensure steady-state conditions. For a more accurate determination of the constants, another point with less than 2 ng/ul of DNA needs to be performed.

Discussion:

Fluorescence molecules provide a non-carcinogenic alternative to radio labeled nucleotides. Use of individually labeled dNTPs allows for the generation of PCR products that are labeled along their entire length. Using a trichloroacetic acid precipitation to remove undigested fragments from solution, the fluorescence of the supernatant can be measured and used to determine the rate of digestion. One drawback of fluorescent labels is their size, relative to including a radio label in the DNA. We proved that, for RecJ, this did not inhibit digestion, but work needs to be performed to show that fluorophores do not change the rate of digestion. This data showed that RecJ has a linear range of digestion of less than 15 min. The linear range of digestion is necessary to know for designing a Michaelis-Menten assay of RecJ digestion rate, where the linear range indicates the time-frame to measure an initial velocity. Interestingly, RecJ digested little more than half of the DNA present in solution, 0.26 µmol digested of 0.48µmol total. This is probably because of enzyme degradation from extended periods of time in high heat.

As another proof-of-concept, we tested the substrate specificity of RecJ, well known to be a 5' to 3' ssDNA exonuclease (Han et al., 2006). Our results were in accordance with this knowledge, showing RecJ to prefer ssDNA to dsDNA substrates. On dsDNA, RecJ retained the greatest efficacy on 5'

overhangs. It was less effective on 3' overhangs and the least effective on dsDNA with blunt ends. Because my substrate was labeled throughout, there is no way to determine which direction RecJ degraded the ssDNA. The results suggestion 5' exonuclease activity. This can be confirmed by blocking either the 5' or the 3' end of ssDNA substrate and measuring the effect this has on the rate of digestion. Blocking the 3' end should have little effect on the rate of digestion while blocking the 5' end should greatly reduce the rate of digestion.

Finally, we assayed the kinetics of RecJ digestion. A Michaelis-Menten analysis of RecJ digestion was used to measure the maximal velocity, the Michaelis constant, and determine the rate of catalysis. V_{max} was found to be 3.6 µM/min and K_m was found to be 36.4 µM. The rate of catalysis, k_{cat}, was determined from the maximal velocity and found to be 3.2 s^{-1}. However, this experiment was performed for times from 10 to 28 minutes, well outside the linear range of digestion as determined from the initial kinetics experiment. To be confident of the values of V_{max} and K_m, this experiment needs to be redone, keeping the time points to less than 15 minutes. These tests have shown that this is an effective method for an exonuclease assay. Next, it will be used to determine the activity of my purified protein and determine if labeling has an effect on the activity of RecJ.

Exonuclease Substrate Design

Fragment Primers		
Forward	gtgcac tggcca gagctc GTGAAACAACAGATACAACTTCGTCGCCG	Tm= 70 C
Reverse (RC)	gcatgcgccggcctaggtgttcgatctcttttcgcgtctggtttagc	Tm= 71 C
5' enzymes		**3' enzymes**
Apa1: gtgcac 5' Msc1: tggcca blunt Sac1: gagctc 3'		AvrII: cctagg 5' Nae1: gccggc blunt Sph1: gcatgc 3'

Figure 1. Exonuclease Construct. Primer design for the exonuclease substrate. Multiple digestion-sites allow for ends of each type (5', 3', double-stranded) to be generated. The digestion-sites added to each end of the fragment are listed at the bottom of the table. They were added in the order that they are listed.

Exonuclease Reactions

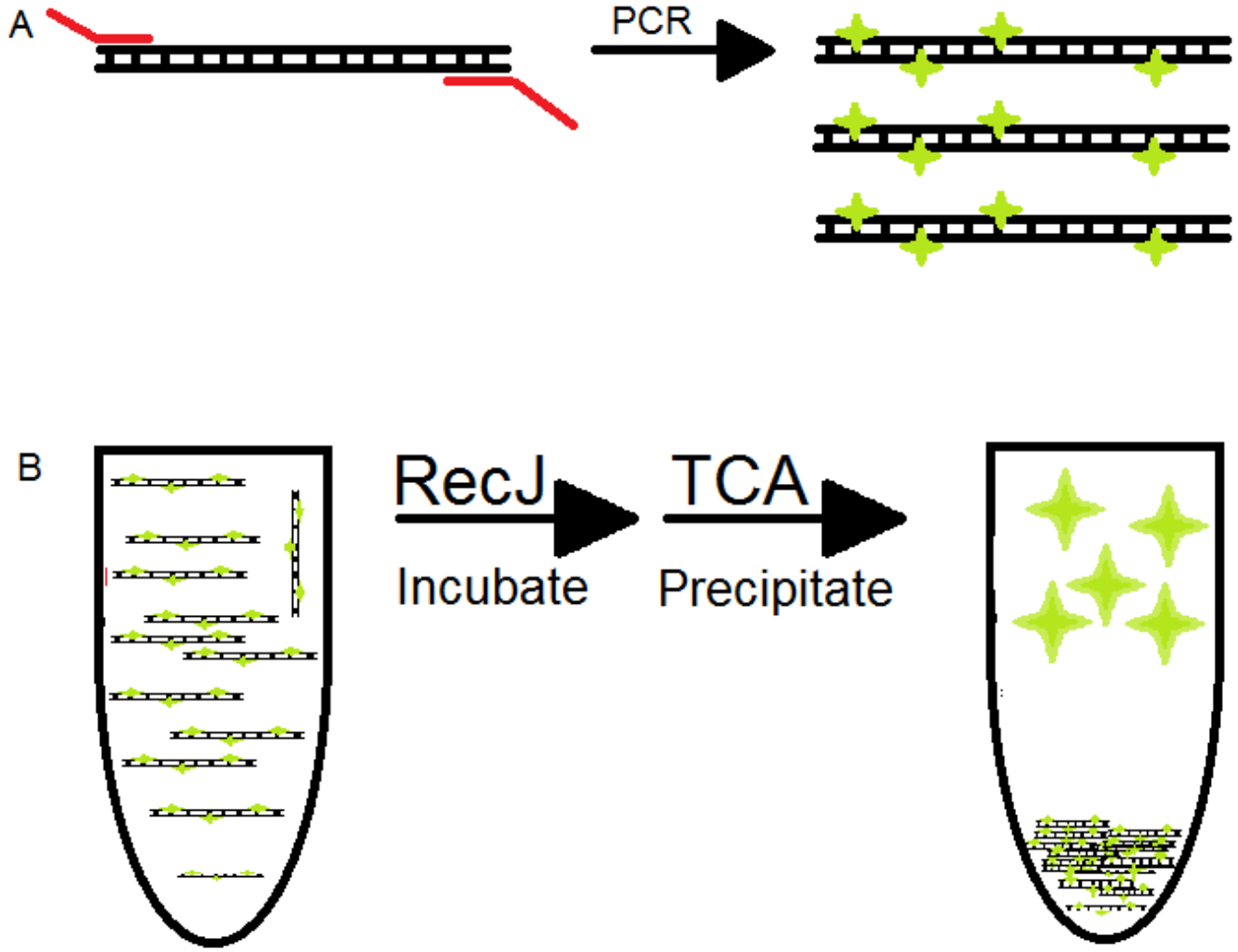

Figure 2. Exonuclease Overview. A) Using primers from figure 1 annealed to pET9a-RecJ, PCR with Cy3-labeled fluorophores was used to generate a 1036bp fragment that is labeled throughout. **B)** Scheme for exonuclease reactions. Labeled DNA was incubated with RecJ. Afterwards, undigested fragments were precipitated using TCA, and free nucleotides were visualized in solution.

Substrate Design

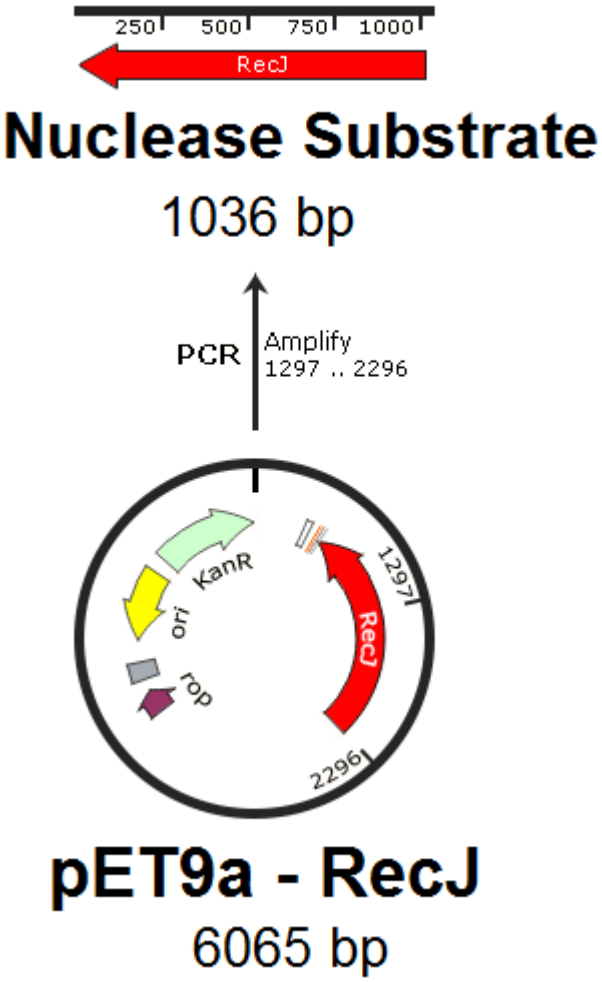

Figure 3. Nuclease Substrate. Final design of exonuclease substrate. PCR with the primers from figure 1 were used to generate a 1036 bp fragment with multiple restriction sites on each end. Fluorophore labeled nucleotides (Cy3-dCTP) were used in the reaction to label the fragment along its entire length. This substrate was used in time-course experiments, specific activity experiments, and determination of V_{max}, K_M and k_{cat} for RecJ.

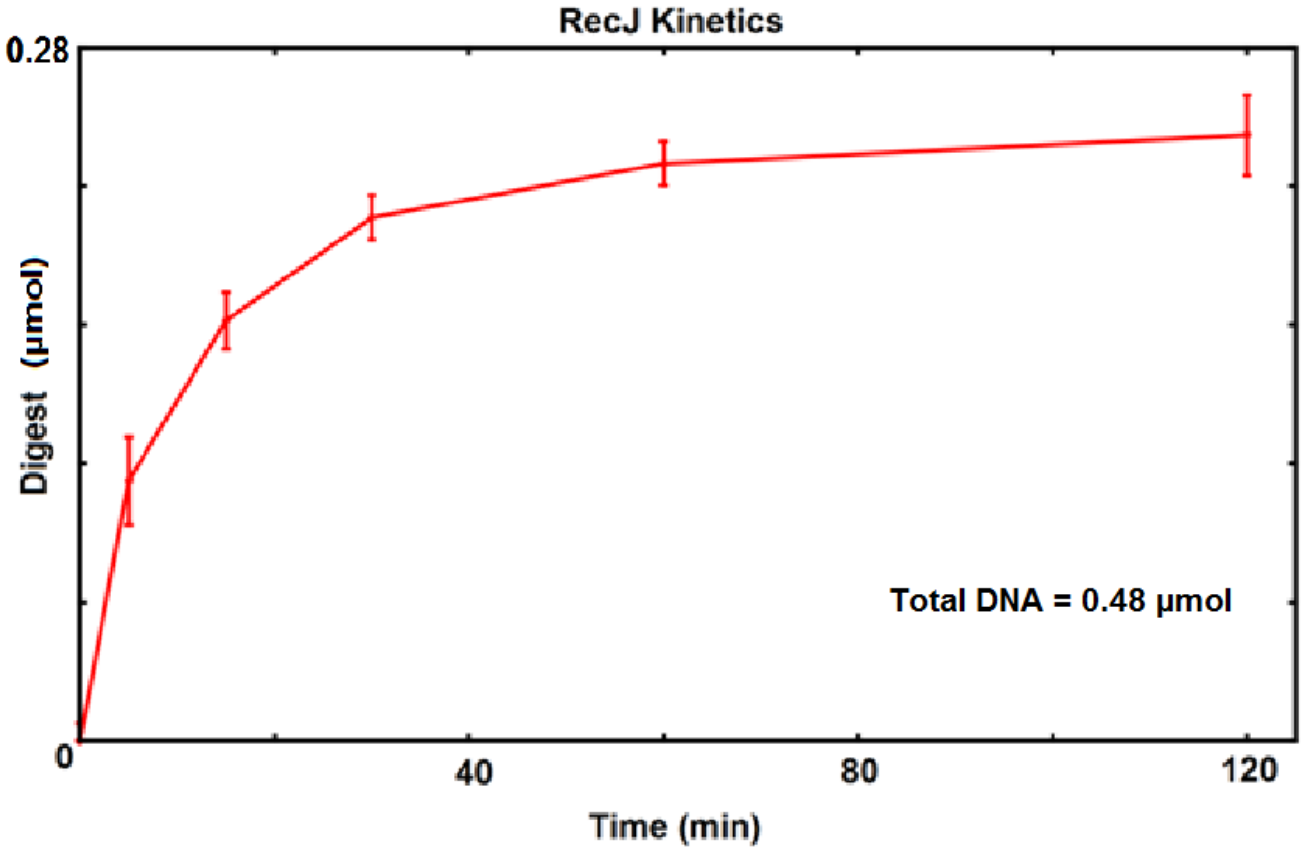

Figure 4. RecJ Kinetics. This was performed with RecJ from NEB to test efficacy of the experimental design. From this, we learned that fluorophore labeled DNA may be digested and visualized using a fluorometer. The linear range of digestion, when we can measure an initial velocity while under steady-state conditions, is less than 20 min.

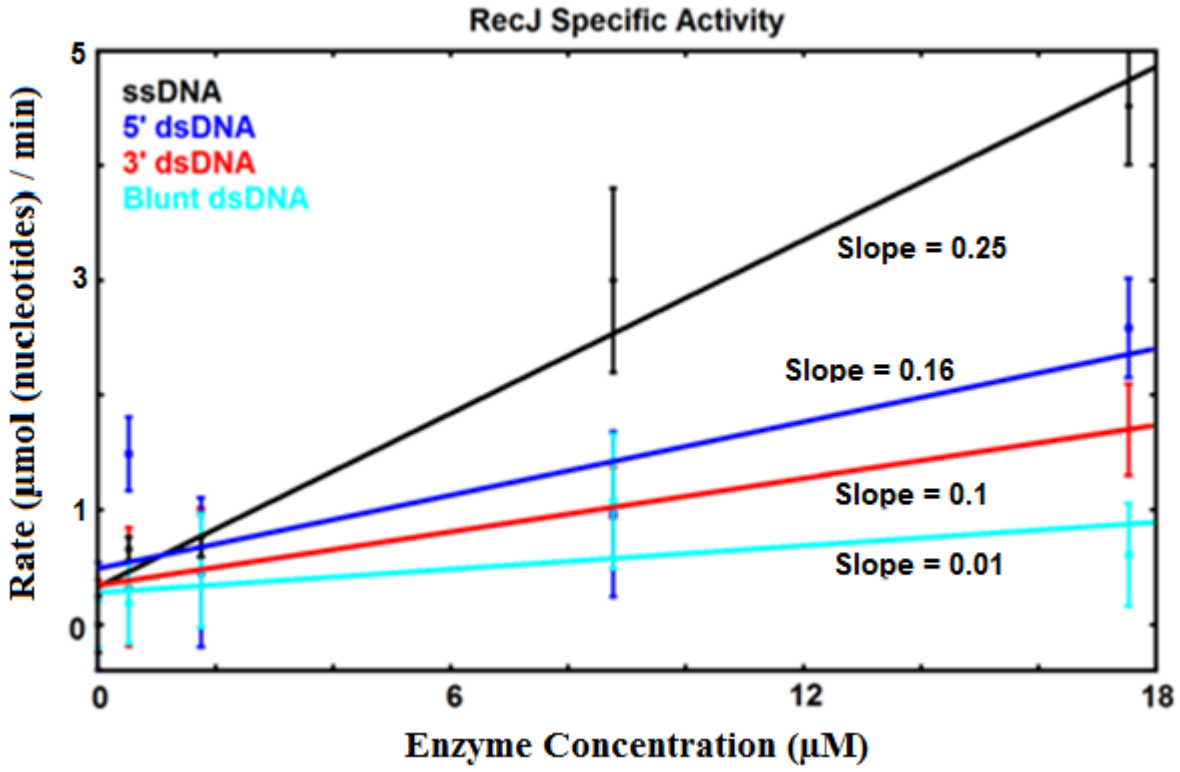

Figure 5. RecJ Substrate Activity. The exonuclease substrate (figure 3) was boiled for 10 min then precipitated on ice, creating single-stranded DNA, digested with ApaI and AvrII, generating 5' overhangs, digested with SacI and SphI, generating 3' overhangs, or undigested to remain as blunt-ended double stranded DNA. These four substrates were then incubated with RecJ to determine its activity on each type. This showed that RecJ is primarily a 5'-3' exonuclease. However, it does retain activity on other substrates with varying levels of efficacy.

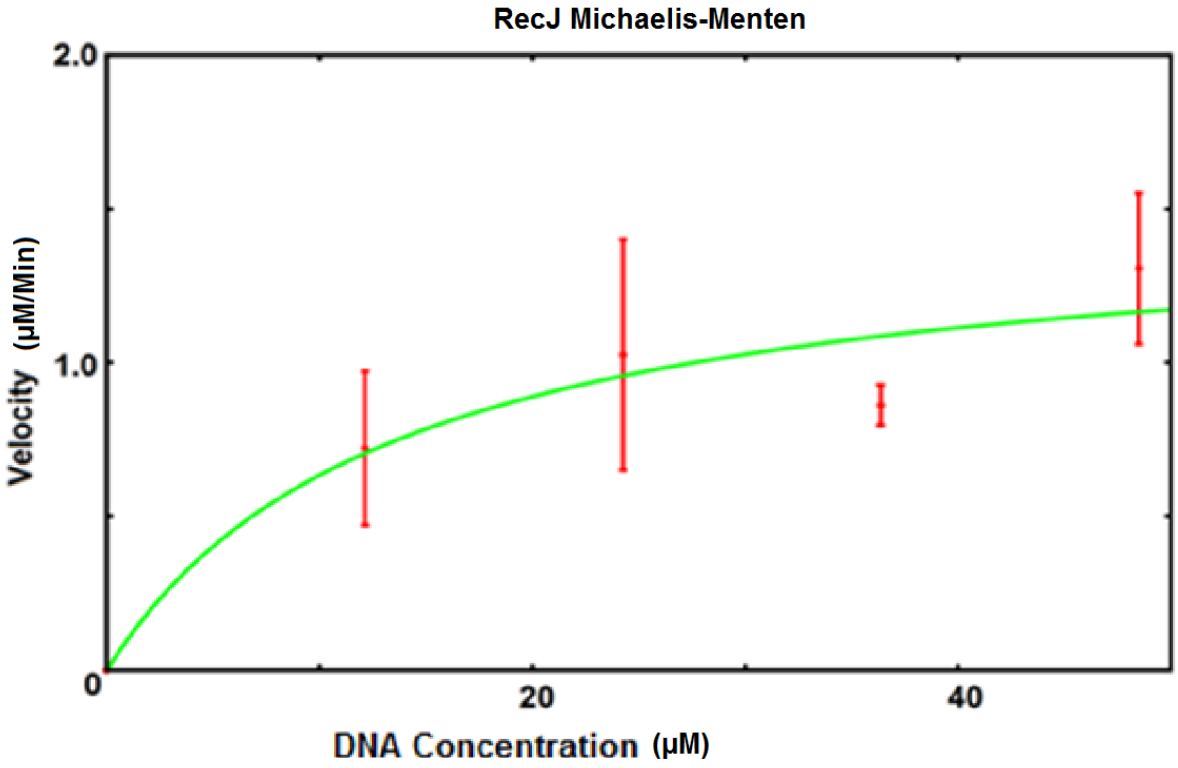

Figure 6. Michaelis-Menten Analysis of DNA Substrate Concentration on RecJ Exonuclease Activity. RecJ was incubated with four concentrations of DNA to determine the effect of DNA concentration on exonuclease activity. The velocity of each reaction is plotted and fit to a Michaelis-Menten curve. From this, we could extract the maximum velocity, V_{max}, and the Michaelis constant, K_M. V_{max} was found to be 3.6 µM/min and K_M was found to be 36.4 µM. Using the V_{max} we determined k_{cat}, the effective rate of the enzyme, to be 3.2 s^{-1}.

Concluding Remarks:

Mechanistic determination of biological processes requires some method for visualizing the protein complexes involved. A short tag that is recognized and converted by FGE provides an efficient and specific method for generating formylglycines on a protein. These forymglycines may then be used in a Hydrazino-Pictet-Spengler (HIPS) reaction to link a high-yield fluorophore to the protein. MutS was the first protein to undergo this method of labeling. It was used to test the efficacy of this process, and we determined that it was an effective way to label proteins. PCNA fusion constructs containing the FGE tag were successfully generated. The next step is to purify and test HIPS-fluorophore labeling of PCNA. We believe that the RecJ protein is the first exonuclease to be fluorophore-labeled *in vitro*. It has been determined that labeling, on the N-terminus, is highly specific and efficient. At the same time, a novel enzyme assay, replacing radio labeled nucleotides with fluorophore-labeled ones, was developed using commercial RecJ. In the future, this assay will be applied to purified RecJ containing a FGE tag and its specific activity compared to the RecJ protein without the FGE-tag will be examined. It is our goal to use the fluorophore-labeled RecJ in Single-Molecule studies.

References:

Acharya, S., Foster, P.L., Brooks, P., and Fishel, R. (2003). The coordinated functions of the E. coli MutS and MutL proteins in mismatch repair. Molecular Cell *12*, 233-246.

Agarwal, P., Kudirka, R., Albers, A.E., Barfield, R.M., de Hart, G.W., Drake, P.M., Jones, L.C., and Rabuka, D. Hydrazino-Pictet-Spengler ligation as a biocompatible method for the generation of stable protein conjugates. Bioconjugate chemistry *24*, 846-851.

Alberts, B., Johnson, A., and Lewis, J. (2002). Molecular Biology of the Cell. , 4 edn (New York, Garland Science).

Axelrod, D., Koppel, D.E., Schlessinger, J., Elson, E., and Webb, W.W. (1976). Mobility measurement by analysis of fluorescence photobleaching recovery kinetics. Biophysical journal *16*, 1055-1069.

Bilezikian, S.B., Nossel, H.L., Butler, V.P., Jr., and Canfield, R.E. (1975). Radioimmunoassay of human fibrinopeptide B and kinetics of fibrinopeptide cleavage by different enzymes. The Journal of clinical investigation *56*, 438-445.

Bornhorst, B.J., and Falke, J.J. (2010). Reprint of: Purification of Proteins Using Polyhistidine Affinity Tags. Protein expression and purification.

Bowman, G.D., O'Donnell, M., and Kuriyan, J. (2004). Structural analysis of a eukaryotic sliding DNA clamp-clamp loader complex. Nature *429*, 724-730.

Carrico, I.S., Carlson, B.L., and Bertozzi, C.R. (2007). Introducing genetically encoded aldehydes into proteins. Nature chemical biology *3*, 321-322.

Cho, W.K., Jeong, C., Kim, D., Chang, M., Song, K.M., Hanne, J., Ban, C., Fishel, R., and Lee, J.B. (2012). ATP alters the diffusion mechanics of MutS on mismatched DNA. Structure *20*, 1264-1274.

Christie, R.J., Anderson, D.J., and Grainger, D.W. Comparison of hydrazone heterobifunctional cross-linking agents for reversible conjugation of thiol-containing chemistry. Bioconjugate chemistry *21*, 1779-1787.

Ciccia, A., and Elledge, S.J. (2011). The DNA damage response: making it safe to play with knives. Mol Cell *40*, 179-204.

Cregg, J.M., Cereghino, J.L., Shi, J., and Higgins, D.R. (2000). Recombinant protein expression in Pichia pastoris. Molecular biotechnology *16*, 23-52.

Fishel, R., Acharya, S., Berardini, M., Bocker, T., Charbonneau, N., Cranston, A., Gradia, S., Guerrette, S., Heinen, C.D., Mazurek, A., *et al.* (2000). Signaling Mismatch Repair: the mechanics of an adenosine-nucleotide molecular switch. Cold Spring Harbor Symp Quant Biol *65*, 217-224.

Forties, R.A., and Wang, M.D. (2014). Discovering the Power of Single Molecules. Cell *157*, 4-7.

Grilley, M., Welsh, K.M., Su, S.S., and Modrich, P. (1989). Isolation and characterization of the Escherichia coli mutL gene product. J Biol Chem *264*, 1000-1004.

Ha, T., and Tinnefeld, P. (2012). Photophysics of fluorescent probes for single-molecule biophysics and super-resolution imaging. Annual review of physical chemistry *63*, 595-617.

Han, E.S., Cooper, D.L., Persky, N.S., Sutera, V.A., Jr., Whitaker, R.D., Montello, M.L., and Lovett, S.T. (2006). RecJ exonuclease: substrates, products and interaction with SSB. Nucleic Acids Res *34*, 1084-1091.

Hengen, P. (1995). Purification of His-Tag fusion proteins from Escherichia coli. Trends Biochem Sci *20*, 285-286.

Jaiswal, J.K., and Simon, S.M. (2004). Potentials and pitfalls of fluorescent quantum dots for biological imaging. Trends Cell Biol *14*, 497-504.

Jeong, C., Cho, W.K., Song, K.M., Cook, C., Yoon, T.Y., Ban, C., Fishel, R., and Lee, J.B. (2011). MutS switches between two fundamentally distinct clamps during mismatch repair. Nat Struct Mol

Biol *18*, 379-385.
Jonsson, Z.O., Hindges, R., and Hubscher, U. (1998). Regulation of DNA replication and repair proteins through interaction with the front side of proliferating cell nuclear antigen. Embo J *17*, 2412-2425.
Kolodner, R.D., and Marsischky, G.T. (1999). Eukaryotic DNA mismatch repair. [Review] [84 refs]. Current Opinion in Genetics & Development *9*, 89-96.
Kolodner, R.D., Mendillo, M.L., and Putnam, C.D. (2007). Coupling distant sites in DNA during DNA mismatch repair. Proc Natl Acad Sci U S A *104*, 12953-12954.
Kost, T.A., Condreay, J.P., and Jarvis, D.L. (2005). Baculovirus as versatile vectors for protein expression in insect and mammalian cells. Nature biotechnology *23*, 567-575.
Lahue, R.S., Au, K.G., and Modrich, P. (1989). DNA mismatch correction in a defined system. Science *245*, 160-164.
Lahue, R.S., Su, S.S., and Modrich, P. (1987). Requirement for d(GATC) sequences in Escherichia coli mutHLS mismatch correction. Proc Natl Acad Sci U S A *84*, 1482-1486.
Lau, P.J., and Kolodner, R.D. (2003). Transfer of the MSH2.MSH6 complex from proliferating cell nuclear antigen to mispaired bases in DNA. J Biol Chem *278*, 14-17.
Marinus, M.G. (1976). Adenine methylation of Okazaki fragments in Escherichia coli. J Bacteriol *128*, 853-854.
Martin-Lopez, J.V., Barrios, Y., Medina-Arana, V., Andujar, M., Lee, S., Gu, L., Li, G.M., Ruschoff, J., Salido, E., and Fishel, R. The hMSH2(M688R) Lynch syndrome mutation may function as a dominant negative. Carcinogenesis *33*, 1647-1654.
Medintz, I.L., Uyeda, H.T., Goldman, E.R., and Mattoussi, H. (2005). Quantum dot bioconjugates for imaging, labelling and sensing. Nature materials *4*, 435-446.
Meinhart, C.D., and Wereley, S.T. (2003). the theory of diffraction-limited resolution in microparticle image velocimetry. Meas Sci technol *14*, 1047-1053.
Miroux, B., and Walker, J.E. (1996). Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. Journal of molecular biology *260*, 289-298.
Modrich, P. (1989). Methyl-directed DNA mismatch correction. J Biol Chem *264*, 6597-6600.
Modrich, P. (1997). Strand-specific mismatch repair in mammalian cells [Review]. Journal of Biological Chemistry *272*, 24727-24730.
Monico, C., Capitanio, M., Belcastro, G., Vanzi, F., and Pavone, F.S. (2013). Optical Methods to Study Protein-DNA Interactions in Vitro and in Living Cells at the Single-Molecule Level. International journal of molecular sciences *14*, 3961-3992.
Nallamsetty, S., and Waugh, D.S. (2007). A generic protocol for the expression and purification of recombinant proteins in Escherichia coli using a combinatorial His6-maltose binding protein fusion tag. Nature protocols *2*, 383-391.
Sharpless, N.E., and Flavin, M. (1966). The reactions of amines and amino acids with maleimides. Structure of the reaction products deduced from infrared and nuclear magnetic resonance spectroscopy. Biochemistry *5*, 2963-2971.
skoog, D.A. (2006). Principles of Instrumental Analysis, 6 edn (Belmont, CA, Thompson Brooks/Cole).
Su, S.-S., and Modrich, P. (1986). Escherichia coli mutS-encoded protein binds to mismatched DNA base pairs. Proc Natl Acad Sci U S A *83*, 5057-5061.
Sutera, V.A., Jr., Han, E.S., Rajman, L.A., and Lovett, S.T. (1999). Mutational analysis of the RecJ exonuclease of Escherichia coli: identification of phosphoesterase motifs. J Bacteriol *181*, 6098-6102.

Terpe, K. (2006). Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. Applied microbiology and biotechnology *72*, 211-222.

Tippin, B., Pham, P., Bransteitter, R., and Goodman, M.F. (2004). Somatic hypermutation: a mutational panacea. Advances in protein chemistry *69*, 307-335.

Uhlen, M. (2008). Affinity as a tool in life science. BioTechniques *44*, 649-654.

Viswanathan, M., and Lovett, S.T. (1998). Single-strand DNA-specific exonucleases in Escherichia coli - roles in repair and mutation avoidance. Genetics *149*, 7-16.

Walker, J.E., Saraste, M., Runswick, M.J., and Gay, N.J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. Embo J *1*, 945-951.

Wang, T.C., and Smith, K.C. (1988). Different effects of recJ and recN mutations on the postreplication repair of UV-damaged DNA in Escherichia coli K-12. J Bacteriol *170*, 2555-2559.

Welsh, K.M., Lu, A.L., Clark, S., and Modrich, P. (1987). Isolation and characterization of the Escherichia coli mutH gene product. J Biol Chem *262*, 15624-15629.